

ORSAY

N° D'ORDRE :

**UNIVERSITÉ DE PARIS-SUD
UFR. SCIENTIFIQUE D'ORSAY**

THÈSE
présentée
Pour obtenir

Le GRADE de DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS XI ORSAY

PAR

Lionel MÉDINI

**SUJET : ACCÈS À L'INFORMATION ORIENTÉ-UTILISATEUR PAR
RECONTEXTUALISATION ET ARTICULATION
DU BESOIN D'INFORMATION.
APPLICATION À LA MAÎTRISE DU RISQUE DE CRITICITÉ.**

Soutenue le 7 septembre 2001 devant la Commission d'examen

M. Michel BEAUDOIN-LAFON	examineur
M. Mathias CHAILLOT	examineur
M. Christian FLUHR	examineur
Mme Jocelyne NANARD :	rapporteur
M. Philippe PALANQUE :	rapporteur
Mme Véronique ROUYER	Invitée
M. Dominique-Louis SCAPIN	directeur

Remerciements

Je tiens à remercier toutes les personnes qui m'ont aidé ou supporté pour la réalisation de ce travail.

Je considère ce travail comme une synthèse des apports des différentes équipes des unités où j'ai été amené à travailler durant ces presque quatre années de thèse, et les remercie pour ces apports.

L'équipe du SEC de l'IPSN, qui m'a réservé un accueil très favorable lors des différentes étapes du projet à partir duquel est bâti cette thèse, et particulièrement Véronique Rouyer pour m'avoir fait découvrir le monde de la sûreté-criticité, et Jacques Anno, pour sa contribution au projet.

L'équipe du projet MErLin de l'INRIA, et tout particulièrement mon directeur de thèse Dominique Scapin, pour la « constructivité » - si je peux me permettre ce premier néologisme du manuscrit – de nos discussions, ainsi que Wendy Mackay, qui m'a fait découvrir les joies de la pluridisciplinarité lors de mon projet de DEA.

Tous les membres de l'équipe de l'ex SIIA de la DIST du CEA – et en particulier Jean-Luc Simoni et Stéphane Picard – pour leurs apports à ce travail et pour tous les autres types d'apports, Mathias Chaillot, pour m'avoir finalement permis de travailler dans les meilleures conditions, ainsi que Christian Fluhr, qui m'a fait l'honneur de participer au jury, et avec qui je regrette de n'avoir pas plus travaillé.

L'équipe du LIA de l'Université d'Avignon, pour m'avoir permis de rédiger ce manuscrit dans les meilleures conditions et m'avoir apporté certaines connaissances en informatique théorique très utiles pour la programmation du prototype, pour le plaisir que j'ai pris à y enseigner durant cette première année d'ATER et pour m'avoir accordé le droit de recommencer.

Je tiens également à remercier Jocelyne Nanard et Philippe Palanque, qui ont accepté de rapporter ce travail, et m'ont apporté des remarques pertinentes à partir desquelles j'espère avoir su améliorer la qualité de ce manuscrit, ainsi que Michel Beaudoin-Lafon, qui m'a fait l'honneur d'accepter de participer au jury de cette thèse, et durant les cours de DEA duquel j'ai pu découvrir le domaine des IHM.

Je remercie également Catherine, pour les heures qu'elle a passées à corriger ce manuscrit, pour m'avoir supporté pendant toute la phase de rédaction, et pour avoir arrêté de fumer.

Je dédicace le présent manuscrit à ma famille, Laurence, Robert, et particulièrement à ma mère Astrid, sans qui...

Table des matières

TABLE DES MATIÈRES	5
CHAPITRE 1. INTRODUCTION	11
BIBLIOGRAPHIE	16
CHAPITRE 2. REVUE DE QUESTIONS : LES DIFFÉRENTS ÉLÉMENTS DE LA PROBLÉMATIQUE DE L'ACCÈS À L'INFORMATION	17
2.1. LE PROBLÈME D'INFORMATION	18
2.2. L'ACCÈS À L'INFORMATION	19
2.2.1. <i>Les stratégies analytiques</i>	<i>20</i>
2.2.2. <i>Systèmes à base d'interrogation</i>	<i>22</i>
2.2.2.1. L'indexation documentaire	23
2.2.2.1.1. Les méthodes d'indexation « classiques »	24
2.2.2.1.2. Vers une approche conceptuelle de l'indexation.....	25
2.2.2.1.3. L'indexation de la requête.....	27
2.2.2.2. La comparaison documentaire	29
2.2.2.3. Les critères d'évaluation	30
2.2.3. <i>Les stratégies de parcours.....</i>	<i>31</i>
2.2.4. <i>Systèmes hypertextes et hypermédia</i>	<i>33</i>
2.2.4.1. Modèles d'hypertextes	34
2.2.4.2. Désorientation et cartes conceptuelles	37
2.2.4.3. Les systèmes hypertextes d'accès à l'information	39
2.2.5. <i>Un point de terminologie.....</i>	<i>39</i>
2.2.6. <i>Spécificités des deux types de stratégies</i>	<i>41</i>
2.2.7. <i>Les stratégies alternées et les systèmes correspondants.....</i>	<i>44</i>
2.2.8. <i>Les stratégies hybrides et les outils associés</i>	<i>47</i>
2.2.9. <i>Bilan de l'accès à l'information.....</i>	<i>48</i>
2.3. DIFFÉRENTES APPROCHES D'AIDE À LA COMPRÉHENSION DE L'INFORMATION....	49
2.3.1. <i>L'approche cognitivienne.....</i>	<i>49</i>
2.3.2. <i>L'approche sémiologique et sémiotique.....</i>	<i>50</i>
2.3.3. <i>La gestion des connaissances.....</i>	<i>52</i>
2.3.4. <i>L'approche ergonomique comme moyen de contrôle des outils d'accès à l'information.....</i>	<i>54</i>
2.3.4.1. Les méthodes de modélisation et d'analyse de tâches	55
2.3.4.2. Les profils d'utilisateurs	56
2.3.4.3. Les règles de conception	57
2.3.4.4. Les méthodes d'évaluation.....	58
2.4. BILAN DE L'ÉTAT DE L'ART DE LA PROBLÉMATIQUE DE L'ACCÈS À L'INFORMATION	61
BIBLIOGRAPHIE	63
CHAPITRE 3. MÉTHODES ET OUTILS : DISCIPLINES IMPLIQUÉES ET POSITIONNEMENT OPÉRATIONNEL DE LA THÈSE PAR	

RAPPORT À LA PROBLÉMATIQUE DE L'ACCÈS À L'INFORMATION 72

3.1. DÉFINITIONS 72

3.1.1.1.	Information.....	72
3.1.1.2.	Information Scientifique et Technique.....	73
3.1.1.3.	Systèmes d'information.....	74
3.1.1.4.	Éléments du SI et granularité de l'information	74
3.1.1.5.	Connaissance	75
3.1.1.6.	Gestion des Connaissances.....	76
3.1.1.7.	Ontologies et types de classification	77
3.1.1.8.	Utilisateur	78

3.2. DISCIPLINES 78

3.2.1. La Gestion des Connaissances..... 78

3.2.1.1.	La méthode de gestion des connaissances MKSM	78
3.2.1.1.1.	Les différents types de méthodes de gestion des connaissances	78
3.2.1.1.2.	Les fondements théoriques de la méthode MKSM	80
3.2.1.1.3.	Les formalismes diagrammatiques de la méthode MKSM.	84
3.2.1.1.3.1.	Le modèle du domaine	84
3.2.1.1.3.2.	Le modèle des activités	86
3.2.1.1.3.3.	Le modèle des tâches.....	87
3.2.1.1.3.4.	Le modèle des concepts.....	89
3.2.1.1.4.	Les évolutions de la méthode MKSM.....	90
3.2.1.1.5.	Méthodologie de modélisation	91
3.2.1.1.6.	Utilisabilité et réutilisabilité de la méthode MKSM	92
3.2.1.1.7.	Opérationnabilité de la méthode MKSM	93
3.2.1.2.	Le Livre de Connaissances Électronique	94
3.2.1.2.1.	Le Livre de Connaissances.....	94
3.2.1.2.2.	Le Livre de Connaissances Hypermédia.....	95
3.2.1.2.3.	Le Livre de Connaissances Électronique	97

3.2.2. L'Ergonomie..... 99

3.2.2.1.	Positionnement et historique de la discipline	100
3.2.2.2.	La méthode de conception ergonomique appliquée à la conception du LMCE	101

3.2.2.2.1.	Conception du Livre de Connaissances Hypermédia.....	101
------------	------------------------------------------------------	-----

3.2.2.2.2.	Conception du système d'accès à l'information.....	102
------------	----------------------------------------------------	-----

3.2.2.2.2.1.	Utilisation de la gestion des connaissances pour une conception justifiée de l'interface	103
--------------	---------------------------------------------------------------------------------------------------	-----

3.2.2.2.2.2.	Modélisation des besoins en information.....	104
--------------	----------------------------------------------	-----

3.2.2.2.2.3.	Modélisation des utilisateurs	105
--------------	-------------------------------------	-----

3.2.2.2.2.4.	Identification des tâches utilisateurs	106
--------------	----------------------------------------------	-----

3.2.2.2.2.5.	Conception de l'interface	106
--------------	---------------------------------	-----

3.2.2.2.2.6.	Évaluation et validation du prototype et de l'approche.....	106
--------------	-------------------------------------------------------------	-----

3.3. LES TECHNOLOGIES EMPLOYÉES POUR LA RÉALISATION DU PROTOTYPE 107

3.3.1. La technologie W3 et XML..... 107

3.3.2. L'approche DVP

3.3.3. La technologie ActiveXâ

3.3.4. Le format pdf et la suite Acrobatâ..... 110

3.4. LE PROJET-PILOTE..... 111

3.4.1. L'unité où a été mené le projet..... 111

3.4.2.	<i>Le domaine de connaissances</i>	112
3.4.3.	<i>La modélisation des connaissances</i>	112
3.4.4.	<i>Inventaire du système d'information</i>	114
3.5.	BILAN DES PERSPECTIVES MÉTHODOLOGIQUES	114
	BIBLIOGRAPHIE	116
 CHAPITRE 4. CONTRIBUTION DE LA THÈSE : L'IDENTIFICATION ET LES ÉLÉMENTS DE SOLUTION PROPOSÉS POUR LA PROBLÉMATIQUE DE L'ACCÈS À L'INFORMATION		
4.1.	PROBLÉMATIQUE	121
4.1.1.	<i>Méthodologie de modélisation</i>	122
4.1.1.1.	Vers la définition d'un modèle de l'utilisateur.....	123
4.1.1.2.	Modélisation du système d'information.....	124
4.1.1.3.	Articulation des modèles pour l'accès à l'information	126
4.1.2.	<i>Vers une stratégie spécifique d'accès à l'information</i>	127
4.1.3.	<i>Problématique opérationnelle</i>	128
4.1.4.	<i>De la pertinence relative de la problématique de l'accès à l'information et de l'approche de la gestion des connaissances</i>	130
4.2.	DÉMARCHE DE CONCEPTION DU LMCE	132
4.2.1.	<i>Les trois dimensions de l'expression du besoin d'information avec le LMCE</i> 133	
4.2.1.1.	Le référentiel diagrammatique du domaine d'activité	133
4.2.1.1.1.	Aspects théoriques.....	133
4.2.1.1.1.1.	Structuration du domaine	133
4.2.1.1.1.2.	Structuration des sous-domaines.....	134
4.2.1.1.1.3.	Structuration des diagrammes	135
4.2.1.1.1.4.	Structuration des éléments des diagrammes.....	135
4.2.1.1.2.	Aspects techniques	136
4.2.1.2.	Les profils d'utilisateurs.....	139
4.2.1.2.1.	Aspects techniques	140
4.2.1.2.2.	Profils par défaut et actualisation des profils	142
4.2.1.2.2.1.	Les profils d'utilisation	142
4.2.1.2.2.2.	Les préférences de visualisation.....	144
4.2.1.2.3.	Méthodologie de modélisation.....	144
4.2.1.3.	Le modèle conceptuel du Système d'Information.....	146
4.2.1.3.1.	Description du modèle du SI.....	146
4.2.1.3.2.	Méthodologie de modélisation	149
4.2.1.3.2.1.	Identification des critères d'indexation	149
4.2.1.3.2.2.	Obtention des index.....	151
4.2.2.	<i>Description fonctionnelle de l'accès à l'information avec le LMCE</i>	151
4.2.2.1.	Le problème d'information comme argument de la fonction d'accès à l'information du LMCE	152
4.2.2.2.	Position des autres éléments de la modélisation	154
4.2.2.3.	La fonction d'accès à l'information.....	154
4.2.3.	<i>Les outils d'expression du besoin d'information</i>	155
4.2.3.1.	La stratégie d'accès à l'information cumulative	155
4.2.3.2.	Le système de règles	157
4.2.4.	<i>Identification des tâches et mise au point des composants de l'interface</i>	159
4.2.4.1.	L'utilisation du LMCE pour la formation	160
4.2.4.1.1.	Les aides à la gestion de la tâche.....	160

4.2.4.1.2.	Les aides à la sélection d'informations	161
4.2.4.1.3.	Les aides à la compréhension.....	164
4.2.4.1.3.1.	Les cartes conceptuelles	164
4.2.4.1.3.2.	Le bandeau de titre	166
4.2.4.2.	L'utilisation du LMCE pour l'accès à l'information	166
4.2.4.2.1.	Description des tâches	167
4.2.4.2.2.	Éléments d'interface.....	170
4.2.4.2.2.1.	Bandeau de titre, diagrammes et cartes conceptuelles	171
4.2.4.2.2.2.	Le multiplexeur de liens	171
4.2.4.2.2.3.	Le caddie à concepts	172
4.2.4.2.2.4.	Les raccourcis.....	176
4.2.4.2.3.	L'interrogation directe.....	177
4.2.4.3.	Les autres fonctions du LMCE.....	177
4.2.4.4.	Les aspects globaux de l'interface	178
4.2.4.4.1.	Navigation dans les diagrammes de connaissances	179
4.2.4.4.2.	Espace titre	180
4.2.4.4.3.	Espace d'échange d'informations	180
4.2.4.4.4.	Espace de commandes utilisateurs	180
4.2.4.4.5.	Accès à l'information utilisant la stratégie cumulative.....	181
4.2.4.4.6.	Accès à l'information par interrogation directe	182
4.2.4.4.7.	Autres fonctions	182
4.2.4.5.	Critères ergonomiques.....	183
4.2.5.	<i>Spécifications du système</i>	188
4.2.5.1.	Fonctionnalités	188
4.2.5.2.	Architecture du LMCE.....	196
4.2.5.3.	Éléments complémentaires des spécifications techniques	197
4.2.5.3.1.	Les bases de données.....	197
4.2.5.3.1.1.	Informations	198
4.2.5.3.1.2.	Méta-données	198
4.2.5.3.2.	Les contenus dynamiques.....	199
4.2.6.	<i>Cycle de vie du LMCE</i>	200
4.2.6.1.	Génération du LMCE	200
4.2.6.2.	Maintenance du LMCE	202
4.3.	APPLICATION AU PROJET-PILOTE	204
4.3.1.	<i>Les modèles conceptuels</i>	204
4.3.1.1.	Le modèle du domaine	204
4.3.1.2.	Le modèle des utilisateurs	205
4.3.1.2.1.	Expert Ingénieur.....	205
4.3.1.2.1.1.	Besoins	205
4.3.1.2.1.2.	Caractéristiques particulières	205
4.3.1.2.1.3.	Tâches.....	206
4.3.1.2.2.	Expert Chercheur.....	206
4.3.1.2.2.1.	besoins	206
4.3.1.2.2.2.	caractéristiques particulières :	207
4.3.1.2.2.3.	Tâches.....	207
4.3.1.2.3.	Apprenant et novice	208
4.3.1.2.3.1.	Besoins	208
4.3.1.2.3.2.	Caractéristiques particulières :	208
4.3.1.2.3.3.	Tâches.....	209
4.3.1.2.4.	Profils d'utilisation	209

4.3.1.3.	Le modèle de l'information.....	211
4.3.2.	<i>Les outils conceptuels.....</i>	212
4.3.3.	<i>Fonctionnalités du système</i>	212
4.3.4.	<i>État d'avancement du prototype.....</i>	212
4.4.	BILAN DE LA CONCEPTION.....	214
	BIBLIOGRAPHIE	216
 CHAPITRE 5. DISCUSSION : L'APPORT ET LES PERSPECTIVES DE NOTRE APPROCHE POUR LA PROBLÉMATIQUE DE L'ACCÈS À L'INFORMATION		
5.1.	SPÉCIFICITÉS DE L'APPROCHE	218
5.1.1.	<i>Aspects méthodologiques</i>	218
5.1.1.1.	Avantages et limites de la démarche de recontextualisation de l'information.....	219
5.1.1.2.	Les stratégies d'accès à l'information proposées par le LMCE	220
5.1.2.	<i>Aspects ergonomiques de notre démarche</i>	222
5.1.2.1.	Présence des caractéristiques du besoin d'information dans les diagrammes	223
5.1.2.2.	Accessibilité des caractéristiques du besoin d'information par la navigation hypermédia dans les diagrammes.....	223
5.1.2.3.	Sélection d'éléments correspondant à des critères d'indexation.....	225
5.1.2.3.1.	Identification des critères de forme	225
5.1.2.3.2.	Identification des critères de fonds.....	226
5.1.2.4.	Sélection des autres éléments des diagrammes	226
5.1.2.4.1.	Éléments de contexte.....	226
5.1.2.4.2.	Éléments de requête	226
5.1.2.4.3.	Intérêt des autres formalismes.....	227
5.1.2.5.	Bilan ergonomique de l'intérêt de la modélisation des connaissances comme moyen de présentation hypermédia des caractéristiques du besoin d'information.....	227
5.1.3.	<i>Apport de l'indexation conceptuelle pour la formalisation du besoin d'information.....</i>	232
5.1.4.	<i>Apports spécifiques à la gestion des connaissances</i>	233
5.1.4.1.	Éléments de réflexion sur la présentation et le positionnement de la méthode de gestion des connaissances MKSM.....	234
5.1.4.2.	Discussion du paradigme d'utilisation de la gestion des connaissances comme référentiel de conception et d'utilisation d'un système d'accès à l'information scientifique et technique (Interactions entre les disciplines pour la conception)	241
5.1.4.2.1.	Les aspects théoriques des modélisations	242
5.1.4.2.2.	Du modèle des activités à la spécification des fonctionnalités du LMCE	243
5.1.4.2.3.	De la modélisation des tâches à la définition de l'interface utilisateur	244
5.1.4.2.4.	Du modèle des concepts à l'indexation sémantique de l'information	246
5.1.4.2.5.	Du modèle du domaine à la création des règles de comportement du LMCE ?	248
5.1.4.2.6.	De l'utilité des autres modèles	249
5.1.4.3.	Pertinence de l'approche de la gestion des connaissances dans le cadre plus général de sa réintroduction dans l'accomplissement de l'activité modélisée ...	250
5.2.	BILAN PROSPECTIF	253
5.2.1.	<i>Processus d'évaluation et de validation.....</i>	254

5.2.1.1.	Inspection ergonomique	255
5.2.1.2.	Tests utilisateurs	258
5.2.1.3.	Questionnaire de satisfaction	260
5.2.1.4.	Analyse des données recueillies et bilan du processus d'évaluation	261
5.2.2.	<i>Perspectives d'évolution de l'outil</i>	262
5.2.2.1.	Perspectives offertes par l'utilisation d'un moteur de règles	262
5.2.2.2.	Vers une indexation conceptuelle multi-supports par le LMCE	263
5.2.3.	<i>Perspectives d'évolution de l'approche d'accès à l'information par recontextualisation</i>	263
BIBLIOGRAPHIE		265
CHAPITRE 6. CONCLUSION		267
BIBLIOGRAPHIE		274
ANNEXES		275
ANNEXE 1. LISTE DES RÉFÉRENCES BIBLIOGRAPHIQUES		276
ANNEXE 2. QUELQUES DIAGRAMMES DE CONNAISSANCES ISSUS DU PROJET-PILOTE		289
ANNEXE 3. FICHIERS XML SOURCES DU PROTOTYPE		299
Annexe 3.1. Fichier XML représentant un diagramme MKSM dans le LMCE		300
Annexe 3.2. Fichier XML représentant un profil d'utilisateur dans le LMCE		314

Chapitre 1. Introduction

Tout individu exerçant une activité à dominante scientifique et technique utilise des informations d'origines et de natures diverses au cours de son travail. Ces informations, extraites d'un *Système d'Information*, lui sont nécessaires pour résoudre des *problèmes d'information* qui se posent dans des situations professionnelles ponctuelles mais fréquentes. L'utilisation de ces informations requiert la recherche de l'information pertinente pour un problème, l'accès à cette information et son interprétation dans un contexte professionnel donné.

Une recherche d'information classique comporte trois phases : formaliser le besoin d'information, effectuer la recherche correspondante et, si plusieurs éléments sont susceptibles de répondre à ce besoin, identifier le(s) plus pertinent(s) d'entre eux. Accéder à l'information consiste éventuellement à la récupérer si elle est sur un support distant et à la « lire » avec un outil adéquat si elle est sur un support informatisé. L'interpréter signifie la comprendre, l'analyser et décider de la réponse qu'elle apporte au problème.

Chacune de ces phases peut générer des difficultés qui perturbent ce processus. La recherche d'information peut ne pas aboutir. En particulier, il est parfois difficile d'exprimer un besoin en information avec une précision et une exactitude suffisante pour générer un ensemble de résultats comprenant l'information recherchée et permettant de l'identifier facilement. L'information peut également être inaccessible ou incomprise. Ceci se produit lorsque l'information est indisponible physiquement ou stockée dans un format ou une langue inconnus de l'individu ou des outils dont il dispose. Enfin, il peut aussi arriver que l'individu ne comprenne que partiellement l'information ou la comprenne mal. Dans ces cas, il pense avoir obtenu la réponse à son problème, et interprète l'information de façon erronée.

Dans tous ces cas, la part de l'activité de l'individu dépendante de cette information s'en trouve perturbée. En effet, si l'information n'est pas retrouvée, est inaccessible ou est entièrement incomprise, la recherche d'information se trouve bloquée. Elle est alors soit remplacée par une solution équivalente plus approximative (raisonnement, intuitions), soit abandonnée car devenue insoluble. Si l'information est mal interprétée, l'individu utilisera des données erronées dans la suite de la séquence des tâches de son travail. Les conséquences de ces phénomènes sont particulièrement sensibles dans les domaines scientifiques et techniques,

où la précision et l'exactitude de l'information sont aussi importantes que l'information elle-même.

Bien entendu, de nombreux types d'outils informatisés permettent d'améliorer et d'automatiser chacune de ces tâches. Ils ont été conçus et développés grâce aux travaux de chercheurs qui se sont intéressés à la résolution de ces problèmes d'information selon différentes approches. Trois d'entre elles sont abordées dans ce document.

L'avènement de l'Internet et des Nouvelles Technologies de l'Information et de la Communication (NTIC) a permis l'essor des Moteurs d'Indexation et de Recherche (MIR) et de la conception d'hypermédias. Ces deux disciplines correspondent à des stratégies de recherche d'information spécifiques et complémentaires. [Agosti & Smeaton 96] présentent un recueil de travaux sur la conception de systèmes d'accès à l'information rapprochant ces deux disciplines.

La Gestion des Connaissances décrit et formalise l'activité humaine. Une méthode de gestion des connaissances comme MKSM ([Ermine & al. 96]) décrit les différentes phases d'un projet qui conduira à la rédaction de *diagrammes de connaissances* explicitant une activité scientifique et technique. Cette discipline ne s'intéresse pas a priori aux problèmes d'accès à l'information, mais permet d'identifier les situations de travail qui conduisent un individu à mener une recherche d'information.

L'Ergonomie des logiciels s'applique entre autres à formaliser les règles de l'interaction entre l'homme et la machine. En particulier, elle fournit des méthodes de mise au point de *profils d'utilisateurs* qui permettent à un système de réagir en fonction des attentes de l'individu qui le manipule. Ces profils facilitent à la fois la recherche d'information et sa restitution en tenant compte des besoins et des préférences de chaque individu (ou utilisateur) pour l'accomplissement d'une tâche donnée.

La complémentarité de ces trois disciplines pour l'accès à l'information est illustrée par les outils de gestion de l'information sous forme de « bases de connaissances » tels que Lotus Notes ou tels que les sites Internet des grands éditeurs de logiciels (MicroSoft, Adobe...). En effet, ces outils de gestion de l'information ont intégré, avec leurs interprétations propres, le concept de connaissance, et se conforment à des règles d'interaction cohérentes.

Cependant, si la prise en compte de l'ergonomie dans la conception logicielle tend à se développer, la gestion des connaissances est encore utilisée le plus souvent pour stocker l'information ou pour y accéder. En effet, un projet de Gestion des Connaissances – tel que

ceux menés à la Direction de l'Information Scientifique et Technique (DIST) du CEA – permet d'identifier toutes les phases de l'activité des différents métiers modélisés. Les situations de recherche d'information y sont donc mentionnées en tant qu'éléments d'une séquence d'activités dans lesquelles apparaissent les flux d'information en circulation. La gestion des connaissances peut donc aider à identifier l'information nécessaire dans un contexte professionnel donné ou à recontextualiser une information par rapport à l'activité qui l'utilise.

D'où l'intérêt d'étudier la validité d'une méthode d'accès à l'information tenant le plus complètement possible compte de l'activité professionnelle dans laquelle cette information est destinée à être utilisée. Cette étude nécessite un outil capable d'identifier les besoins en information des différents utilisateurs puis de les guider dans leurs recherches d'information grâce à la connaissance des types d'information requis par leurs activités. Une telle démarche est pluridisciplinaire et fait appel à des méthodes existantes dans les domaines de la modélisation des connaissances, de l'ingénierie de l'information et de l'ergonomie.

L'originalité de cette approche est qu'elle ne se limite pas à juxtaposer ces trois disciplines mais tient compte des interrelations entre celles-ci. Elle s'intéresse notamment aux limites de la transcription des connaissances sous forme d'information et de la modélisation des connaissances pour l'analyse ergonomique. L'objectif est d'aboutir à une méthode de mise en place d'un tel outil intégrant ces trois disciplines en une seule phase de modélisation.

Le champ de cette étude reste limité aux domaines scientifiques et techniques suffisamment complexes pour justifier une action de capitalisation des connaissances qu'ils mettent en œuvre. En outre, elle n'est intéressante que si l'activité nécessite en soi des accès à l'information suffisamment fréquents et diversifiés. Il existe toutefois d'autres catégories de domaines où une telle problématique pourrait s'appliquer (droit, commerce, sciences humaines...). Cette étude se limite volontairement aux domaines scientifiques et techniques, où la mise en relation de la gestion des connaissances et de l'information nous paraît plus « directe » que dans d'autres types de domaines¹. Il n'a d'autre part pas été démontré qu'une modélisation des connaissances de type MKSM pouvait être effectuée sur des domaines de types plus généraux.

¹ En sciences humaines et sociales par exemple, la recherche d'information met en jeu des processus linguistiques et des inférences sémantiques beaucoup plus complexes que dans les domaines scientifiques et techniques, où le vocabulaire est en général préalablement défini.

L'approche est fondée sur la mise en œuvre d'un projet-pilote opérationnel. Ce projet donne matière à la conception d'une maquette destinée à permettre de dégager l'architecture générique d'un système d'accès à l'information centré sur l'utilisateur et guidé par les modèles de connaissances. Ce système s'appuie sur des modèles conceptuels issus des trois disciplines concernées et propose une stratégie d'accès à l'information spécifique. La spécification de ce système permet ensuite d'initier une discussion sur la complémentarité des méthodes de modélisation dans les trois disciplines abordées. D'autre part, une étude de validation de la maquette est prévue pour apprécier l'intérêt de cette démarche par rapport à d'autres modes d'accès à l'information.

Ce document propose tout d'abord une **revue de question** (chapitre 2), destinée à situer notre problématique en fonction des différentes disciplines abordées. Chacune de ces disciplines apporte un éclairage particulier et des éléments de réponse à la question « Quels sont les facteurs qui sont susceptibles de favoriser la transmission de l'information entre les individus ? » Le propos s'articule autour des différentes étapes de résolution de ce qui, selon [Marchionini 95], est à l'origine d'une démarche d'accès à l'information : le « problème d'information ».

Le chapitre 3 s'intitule « **Méthodes et outils** ». Il s'articule en trois parties. Il commence par proposer des définitions pour les principales notions auxquelles cette thèse fait référence, comme l'information, la connaissance ou l'utilisateur. Puis il positionne les méthodes utilisées dans chacune des différentes disciplines, dont certaines sont définies spécifiquement pour cette thèse, ainsi que les choix technologiques faits pour les développements. Dans un troisième temps, il décrit le projet-pilote sur lequel s'appuient le travail de conception et la réflexion méthodologique qui composent ce travail.

Le chapitre 4 présente **la contribution de la thèse**. Il s'articule en deux parties. Tout d'abord, les différents aspects de la problématique de la thèse sont détaillés. Il s'agit d'identifier une problématique liée à la conception de l'outil, à une méthode de mise au point optimale, du point de vue de la réutilisabilité des composants et des modèles et enfin à l'évaluation de la pertinence de l'approche proposée.

Dans la seconde section de ce chapitre, nous décrivons le travail réalisé pendant la thèse. Il s'agit de donner les descriptions conceptuelle et technique des différents aspects du système d'accès à l'information fondé sur la navigation dans les diagrammes de connaissances, ainsi qu'une description de la stratégie d'accès à l'information qu'il propose. Le prototype sert de point de départ pour les spécifications génériques de l'outil. Le niveau d'avancement des

développements est indiqué pour chaque élément de ce système. Nous présentons en parallèle la méthode de modélisation qui permet d'aboutir aux modèles et aux outils conceptuels nécessaires à la réalisation technique. Le propos est articulé autour de la démarche de conception ergonomique qui a permis d'aboutir à ces spécifications.

Enfin, **la discussion** présentée au chapitre 5 aborde les apports et les limites du travail réalisé. Tout d'abord présentées du point de vue global de la démarche pluridisciplinaire d'accès à l'information, nous discutons ensuite les spécificités ergonomiques de la démarche d'accès à l'information par recontextualisation sous un angle ergonomique. Du point de vue de la gestion des connaissances, nous présentons ensuite une réflexion sur le statut des modèles conceptuels dans la méthode MKSM dans la perspective spécifique de l'accès à l'information. Il en résulte une proposition de réorganisation de ces modèles selon un modèle plus proche des préoccupations des individus à qui est destinée la modélisation. La pertinence de cette approche d'opérationnalisation de la méthode MKSM est alors discutée dans un cadre plus global.

La seconde partie de ce chapitre présente les perspectives qu'offre ce travail. À court terme, la phase d'évaluation du prototype et de l'approche d'accès à l'information par recontextualisation est détaillée. À moyen et à long terme, nous proposons des perspectives techniques d'évolution de l'outil réalisé ainsi que des directions de recherche que nous considérons comme intéressantes à explorer à la suite de ce travail.

Nous concluons en présentant la synthèse de ces travaux et des avancées dans la définition de l'approche proposée. Nous mettons l'accent sur le caractère pluridisciplinaire de la démarche et sur les perspectives que cette pluridisciplinarité permet d'envisager pour l'amélioration des méthodes et des outils d'accès à l'information.

En annexe sont présentés la liste des références bibliographiques tous chapitres confondus, quelques diagrammes de connaissances et des éléments du code du prototype. Les diagrammes de connaissances permettent à la fois de visualiser les formalismes utilisés et de comprendre leur articulation. Ils aident de plus à situer le domaine abordé dans le projet-pilote. Les éléments de code illustrent la structure des modèles conceptuels utilisés dans l'outil. Ils montrent également la séparation qui existe entre les données et les structures de données, ce qui favorise la réutilisabilité de cette structure pour plusieurs projets.

Bibliographie

- [Agosti & Smeaton 96] : *"Information Retrieval and Hypertext"*, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, 1996.
- [Ermine & al. 96] : Ermine J-L., M. Chaillot, P. Bignon, B. Charreton, D. Malavieille: *"MKSM, a method for knowledge systems management"*. in Knowledge Management, Organization, Competence and Methodology, Advances in Knowledge Management, Volume 1, Jos F. Schreinmakers Ed., pp. 288-302, Ergon Verlag, 1996.
- [Marchionini 95] : Marchionini G., *"Information seeking in electronic environments"*, New York, Cambridge University Press, 1995.

Chapitre 2. Revue de questions : les différents éléments de la problématique de l'accès à l'information

La problématique de cette thèse concerne la méthodologie de conception d'un système d'accès à l'information fondé sur une démarche d'accès à l'information pertinente en fonction des besoins de l'utilisateur. Cette problématique est développée au chapitre 4. Pour cela, nous avons besoin de la situer par rapport à la problématique générale de l'accès à l'information. Cette problématique est multidisciplinaire et concerne aussi bien la conception de systèmes de recherche d'information ou d'hypermédias que la gestion des connaissances ou l'ergonomie. Il s'agit dans ce chapitre de positionner la thèse par rapport aux nombreux travaux qui ont déjà été effectués dans ces domaines ou dans des domaines connexes.

Ce chapitre présente un état de l'art de cette problématique en fonction des différentes disciplines concernées. Cet état de l'art n'est pas présenté en fonction de chacune de ces disciplines, mais s'articule autour d'un raisonnement qui séquentialise les différentes étapes rencontrées lors de la démarche d'accès à l'information. Dans chacune de ces étapes viennent s'insérer les éléments des disciplines qui vont nous permettre d'avancer dans la compréhension de l'état de l'art autour de notre problématique.

Bien entendu, nombre de ces disciplines sont vastes et traitent également de problèmes qui ne font pas partie de nos préoccupations. C'est pourquoi cet état n'a pas pour objectif de donner une description de l'intégralité de chacune de ces disciplines, mais se veut pluridisciplinaire en essayant de mettre en lumière l'apport de chacune d'elles pour notre problématique et les relations entre elles.

Ce chapitre commence par la définition du déclencheur de toute démarche d'accès à l'information : le problème d'information. Selon les niveaux de perception de ce problème et l'étendue du manque qu'il met au jour, nous abordons ensuite les différentes stratégies qui peuvent être utilisées par un individu pour accéder à l'information et les différents types d'outils s'y rapportant. Nous nous intéressons ensuite à la façon dont l'individu, désormais en contact avec l'information, l'utilise par rapport à un problème plus général, posé par exemple dans le cadre d'une activité professionnelle à dominante scientifique et technique.

2.1. Le problème d'information

À l'origine de toute démarche d'accès à l'information se situe un manque, un « trou dans le système de connaissances d'un individu »². Pour combler ce manque, l'individu n'a à sa disposition que la solution de rechercher une information lui permettant, par le biais de processus d'interprétation et de compréhension complexes, de recouvrir ce champ d'incomplétude³. C'est pourquoi il est de coutume de qualifier ce manque cognitif de *problème d'information*. Pour expliquer la formation du problème d'information, on se réfère également souvent à [Mackay 60] via [Ingwersen 94] : « a certain incompleteness in his [the inquirer's] picture of the world, –an inadequacy in what we may call his 'state of readiness' to interact purposefully with the world around him... » C'est donc une incomplétude, un manque ou une anomalie qui est déclencheur de processus cognitifs en rapport avec le développement d'un problème d'information.

[Taylor 62] via [Marchionini 95](p. 35) définit quatre niveaux de problème d'information : ressenti⁴, conscient, exprimable et formalisé. Le niveau ressenti est la reconnaissance d'un certain manque, mais qui n'est pas cognitivement défini. Au niveau conscient, l'utilisateur peut le circonscrire, mais n'est pas capable de l'exprimer de manière précise. Au niveau exprimable, il est capable d'explicitement clairement le problème (en langage naturel par exemple). Au niveau formalisé, il est capable de « traduire » ce problème en une description formelle comme le langage de requêtes d'un moteur de recherche ou d'un SGBD.

Pour Gary Marchionini, entre les niveaux ressenti et conscient se situe le « gap » de [Dervin & Nilan 86]. Ces deux niveaux constituent ce que [Belkin 80] appelle un « état anormal de connaissance »⁵. Le passage du niveau conscient au niveau exprimable fait ensuite intervenir le langage puis des inférences logiques pour le niveau formalisé.

Pour être exhaustif, il nous semble cependant qu'il faudrait également considérer un cinquième niveau de problème d'information, encore plus vague que le niveau ressenti, et que nous pourrions qualifier d'inconscient. Ce niveau nous permet alors de distinguer la notion de problème d'information de celle de gap. En effet, à ce niveau, il existe chez l'individu une certaine incomplétude dans son état de connaissance, mais le gap n'est pas atteint. L'individu,

² Traduction libre de ce que G. Marchionini décrit comme « a lack in the person's knowledge » ([Marchionini 95]).

³ Ces processus seront brièvement abordés au paragraphe 2.4.

⁴ Traduction libre du terme « visceral ».

⁵ « Anomalous state of knowledge ».

qui n'a pas conscience de ce problème d'information ne peut alors initier une recherche d'information. La littérature en sciences de l'information fait assez peu de cas de ce niveau de problème d'information, sans doute parce qu'il ne mène à aucune démarche de recherche d'information. En sciences cognitives, et notamment lors de l'observation et l'étude des phénomènes d'apprentissage, des auteurs ont cependant noté des différences de comportement et d'efficacité sensibles entre des individus se situant de part et d'autre de ce niveau de problème d'information ([Schneiderman 76], [Bateson 87] via [Mayer 97], [Allen 97]).

Cette distinction des situations d'accès à l'information en fonction des différents niveaux de perception d'un problème d'information peut paraître peu précise. Notamment, les travaux concernant l'étude des processus cognitifs mis en cause ([Rouet et Tricot 98]) ou la modélisation des tâches d'accès à l'information ([Smeaton 96], [Tricot & Nanard 98]) peuvent être considérés comme plus aboutis. Certains d'entre eux sont d'ailleurs repris plus loin dans cet état de l'art. Cependant, ces travaux sont en général axés sur des applications ou des contextes bien précis, qu'il est difficile de généraliser. Nous faisons référence à ces travaux spécifiques lorsque les conditions d'application le permettent. La notion de problème d'information nous fournit en revanche un cadre structurant générique que nous utilisons dans la suite de ce travail, indépendamment des situations diverses dans lesquelles l'accès à l'information est considéré.

Nous ne considérons ici que les problèmes d'information à partir du niveau conscient car seuls ceux-ci donnent lieu à une démarche de recherche d'information. Le fait d'aider les individus à prendre conscience de leur problème d'information n'est pas du ressort de la problématique d'accès à l'information qui nous occupe, mais plutôt de celle de la compréhension des phénomènes d'apprentissage, qui ne sont pas abordés dans cette thèse.

2.2. L'accès à l'information

À partir du niveau conscient, un problème d'information provoque donc chez un individu une impression de manque (« conscious need », [Taylor 62]), qu'il peut essayer de combler en employant une démarche d'accès à l'information adaptée à ce problème⁶. Dans ce cas, l'individu est confronté à une problématique d'accès à l'information, ou *Information Retrieval* (voir paragraphe 2.2.5 pour des détails sur la traduction de ce terme). Il doit alors choisir ou élaborer une *stratégie d'accès à l'information* qui lui permettra de mener à bien la *tâche*

*d'accès à l'information*⁷. Cette stratégie est choisie et dirigée en fonction d'un but, qui peut lui-même évoluer avec les résultats obtenus. La précision de ce but est évidemment fortement corrélée avec le niveau de perception du problème d'information. Ces stratégies sont difficilement séparables des techniques qui les mettent en œuvre. Pour chacune d'elles, il existe une classe spécifique d'outils d'accès à l'information.

Dans cette partie, nous ne nous intéressons qu'aux informations textuelles. Les autres types d'information – multimédia (images, sons, vidéo) ou de type flux de données (par exemple les fichiers « spool » échangés par les codes de calcul) n'ont pas été traités dans cette thèse. Les paragraphes qui suivent s'intéressent donc aux systèmes d'accès à l'information textuelle, et en décrivent en particulier les méthodes d'indexation, de comparaison ou de structuration. Un tel travail pour d'autres types d'informations nous aurait conduit à aborder des techniques spécifiques comme l'extraction de descripteurs d'images ou non encore stabilisées comme les méthodes d'indexation sonores qui n'apportent rien à la compréhension du travail présenté dans la suite de ce manuscrit.

Les systèmes d'information évoqués ici sont donc constitués de bases contenant des *corpus documentaires*, dont l'élément unitaire est le document. Ceux-ci peuvent être découpés en entités plus petites, comme la page, le paragraphe ou le mot.

Deux grandes catégories de stratégies existent pour accéder à l'information : les stratégies analytiques et de parcours. Ces stratégies sont présentées dans les paragraphes suivant, ainsi que les outils d'accès à l'information qui s'appuient sur ces stratégies. Les spécificités de ces stratégies sont ensuite discutées, et une revue des systèmes proposant des stratégies hybrides est présentée.

2.2.1. Les stratégies analytiques

Un individu confronté à un problème d'information va tenter de l'exprimer sous forme de question. C'est le paradigme question/réponse de [Lelu 93]. Il correspond à l'emploi d'une stratégie dite analytique ou d'interrogation. Demander oralement une information à quelqu'un,

⁶ Voir chapitre des définitions pour la notion d'information. L'information pertinente par rapport à un problème d'information est supposée être celle qui comble au moins exactement le manque laissé par le problème dans l'organisation des connaissances de l'individu.

⁷ La distinction entre stratégie et tâche oppose ici les niveaux cognitif et opérationnel. La stratégie regroupe l'ensemble des actions que l'individu décide de faire pour arriver à un but, tandis que la tâche décrit l'ensemble des actions effectivement réalisées. En outre, une divergence peut exister entre les actions initialement planifiées dans la stratégie et celles de la tâche effectivement réalisée.

interroger une base de données ou un moteur de recherche sur le web sont des emplois courants de ce type de stratégie.

Ce type de stratégie nécessite la formulation aussi précise que possible d'un *besoin d'information*. Celui-ci se caractérise par une description plus ou moins formelle de l'information recherchée. C'est dans la précision de cette formalisation qu'intervient le niveau de précision du problème d'information. En effet, un problème d'information perçu au niveau conscient sera plus dur à formaliser que s'il l'était au niveau exprimable ou formalisé. Besoin et problème d'information se situent donc respectivement à des niveaux informationnel et cognitif.

Pour [Kolmayer 98], la démarche d'accès à l'information par interrogation est loin d'être linéaire. Cet auteur propose en effet plusieurs représentations de la tâche d'information retrieval, dont celle de [Bates 89], représentée ci-dessous.

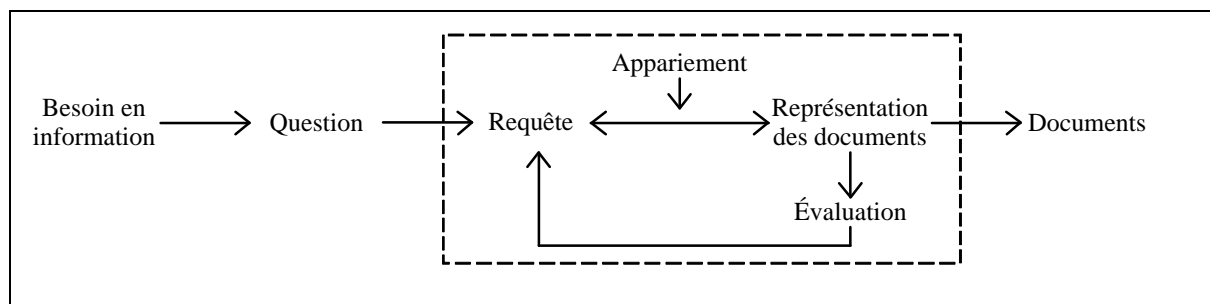


Figure 1. L'accès à l'information par interrogation pour [Bates 89], via [Kolmayer 98].

Ce schéma présente plutôt un aspect fonctionnel des systèmes d'accès à l'information à base d'interrogation. En reprenant ce schéma, Elizabeth Kolmayer précise qu'elle y a rajouté une boucle de feedback, qu'elle a située entre la représentation des documents et la requête. Cette boucle met alors en évidence le caractère non linéaire de ce processus.

Il nous semble cependant que l'idée de Marcia Bates n'était à l'origine que de décrire une seule itération dans une tâche d'accès à l'information par interrogation. Le problème d'information durant l'accomplissement de cette tâche est donc supposé constant, et ce n'est qu'après la phase de représentation des documents que peut s'opérer ce feedback. Cependant, celui-ci n'est pas le fait du système mais de l'utilisateur. En ce sens, il nous semble plus judicieux de « sortir » ce feedback du cadre représentant – d'après la compréhension que nous en avons – le traitement de la requête par le système et de rediriger cette boucle non vers une reformulation de la requête, mais du besoin d'information. Une représentation de ce processus qui nous semble plus appropriée est donnée ci-dessous.

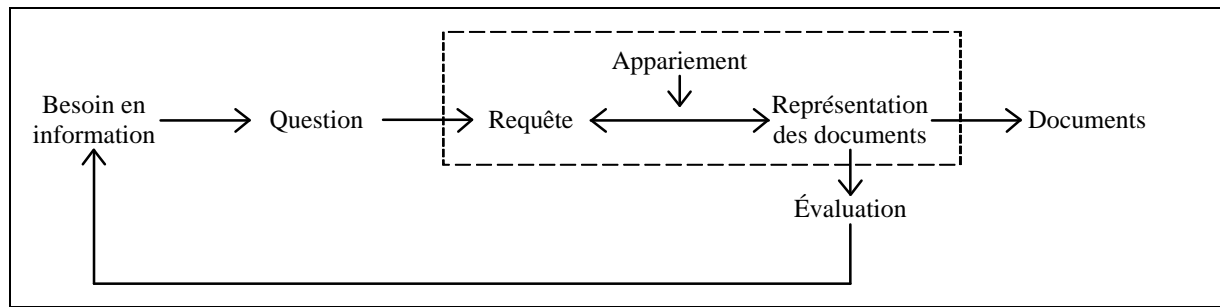


Figure 2. L'accès à l'information par interrogation avec un feedback étendu.

Pour Gary Marchionini (toujours d'après E. Kolmayer), ce feedback aboutit à un repositionnement du problème d'information. Il propose d'ailleurs une représentation de ce processus centrée sur l'utilisateur et sur son problème d'information, comme illustré ci-dessous. Dans ce diagramme, l'auteur met en avant l'aspect non linéaire de la recherche d'information par interrogation pour l'utilisateur.

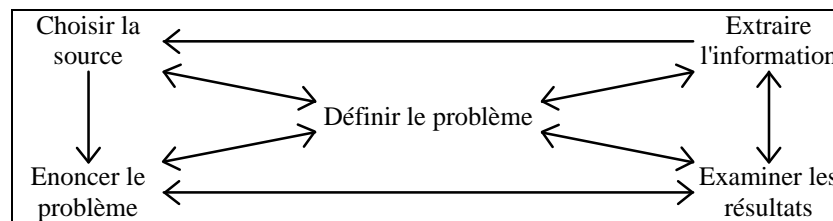


Figure 3. La définition de la recherche d'information par Marchionini.

2.2.2. Systèmes à base d'interrogation

Généralement, les systèmes d'accès à l'information à base d'interrogation, que nous assimilons ici aux systèmes de type *moteurs d'indexation et de recherche*, fonctionnent en essayant de retrouver dans les documents d'une base, des termes de la question ou des termes s'en rapprochant. Grâce à des techniques plus ou moins évoluées⁸, les concepteurs de ces systèmes vont s'efforcer de les rendre à la fois les plus simples d'utilisation et les plus performants possibles⁹. [Leloup 98] présente un panorama de l'offre disponible dans ce domaine. Tous ces systèmes restent toutefois bâtis sur le même modèle, que [Simoni 00] qualifie de « classique en recherche d'information », puisqu'il est employé depuis les années 60. Ce modèle est illustré dans la figure ci-dessous.

⁸ Troncature, reformulation monolingue ou multilingue, analyse morphologique ou grammaticale des termes, définition d'un langage d'interrogation ou interrogation en langage naturel...

⁹ Les performances d'un moteur d'indexation et de recherche se mesurent grâce à plusieurs paramètres, qui sont soit le bruit et le silence, soit le rappel et la précision ([Salton 83]).

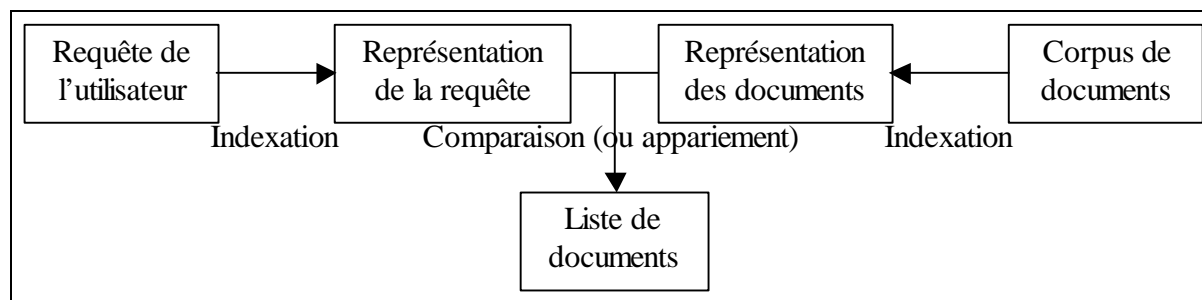


Figure 4. Schéma de fonctionnement d'un système de recherche d'information

Sur cette figure, nous voyons que ce modèle est symétrique : un système d'accès à l'information à base d'interrogation utilise des représentations de la requête et du corpus documentaire réalisées grâce à des *fonctions d'indexation*. Ces fonctions ne sont cependant pas nécessairement identiques du fait de la dissymétrie entre la taille et la structure des entités auxquelles elles s'appliquent : une requête peut être une simple juxtaposition de termes séparés par des opérateurs booléens, tandis qu'un document peut être un article scientifique de plusieurs pages. Bien entendu, l'indexation du corpus est un processus lourd qui n'est réalisé qu'à chaque mise à jour du corpus. L'indexation de la requête doit, elle, se faire à la volée. La comparaison (ou l'alignement) de ces deux représentations est réalisée par une *fonction de comparaison* et a pour but d'obtenir une liste des documents du corpus pertinents pour la requête. Les deux paragraphes suivants, qui présentent succinctement les domaines de l'indexation et de la comparaison documentaires, sont en majeure partie issus de la synthèse de ces domaines présentée dans [Roussey & al. 99].

2.2.2.1. L'indexation documentaire

Le but de l'indexation documentaire est donc d'obtenir une représentation d'un élément textuel. L'indexation commence par l'identification des mots à partir desquels une recherche documentaire est significative ou *descripteurs* des documents. L'approche classique de l'indexation est de définir un jeu de descripteurs pour chaque document. Signalons toutefois l'existence d'une approche fondée sur le morcellement des unités documentaires visant à établir une indexation séparée des parties (ou fragments) des documents. Une telle approche est appelée segmentation documentaire ou « document clustering » (cf. [Lainé-Cruzel & al. 96], [Shapira & al. 96]).

Plusieurs types de représentations sont couramment utilisés. Pour ([Le Loarer 1994], via [Roussey & al. 1999]), quatre méthodes sont principalement utilisées en indexation documentaire « classique ». Celles-ci sont présentées dans la première des deux sous-parties

de ce paragraphe. Dans la seconde, nous abordons la définition et la présentation des travaux réalisés jusqu'à présent dans le domaine de l'indexation conceptuelle. Ce domaine, s'il n'est pas nécessairement orthogonal aux méthodes présentées dans la première sous-partie, a récemment émergé comme une problématique forte de l'indexation documentaire, au point qu'un numéro spécial de la revue « Document Numérique » lui a été consacré ([Jolion 01]).

2.2.2.1.1. *Les méthodes d'indexation « classiques »*

L'indexation à plat est une indexation simple de tous les mots « non vides¹⁰ » du corpus documentaire. L'index réalisé est une liste des mots et des coordonnées de leurs occurrences dans les textes. C'est ce type d'indexation qui est utilisé entre autres par les moteurs de recherche sur le web. À titre d'exemple, d'après J-P. Blanger, directeur de la société BVA MYFRA¹¹ ([Blanger 98]), pour une indexation à plat, en texte intégral simple, comme celle qui est effectuée pour le moteur d'indexation et de recherche Topic de Verity¹², la taille de l'index généré pour un document en Français « libre¹³ » est d'un tiers de celle du document. Pour l'indexation d'un corpus entier, le rapport des tailles de l'index et du corpus est variable en fonction de la cohérence des documents qui le composent (c'est-à-dire de la quantité de documents contenant les mêmes descripteurs).

L'indexation pondérée consiste, en plus d'une indexation à plat, à attribuer des poids aux termes indexés, qui seront utilisés par la fonction de comparaison. Ces poids peuvent être calculés à l'aide de différentes formules. Le critère le plus couramment utilisé dans ces formules est le nombre d'occurrences des mots dans les documents.

L'indexation à rôles permet de relier chaque syntagme nominal d'un élément textuel à un concept du domaine associé. Cette mise en relation est l'assignation d'un rôle à ce nom ou groupe nominal. Pour [Roussey & al. 1999], ce type d'indexation nécessite des « entités d'indexation » pour le domaine et « Seul un système à base de connaissances, propre à un domaine précis, permet de générer ce genre d'index. La lourdeur et la complexité de la

¹⁰ Un mot vide est un mot « correspondant aux catégories lexicales telles qu'adverbes, pronoms etc. [...] jugés inutiles pour l'indexation. » [Rousset & al. 1999].

¹¹ BVA MYFRA commercialise des systèmes de Gestion Électronique de Documents. Cette société est, entre autres, le principal distributeur pour la France des scanners Ricoh, de l'outil de GED Alchemy d'IMR et de plug-ins de GED pour Adobe Acrobat. Elle organise également des formations pour les utilisateurs de ces produits.

¹² Livré gratuitement avec le logiciel Acrobat Reader d'Adobe depuis la version 3.01. L'indexation se fait à partir du logiciel Catalog de la suite payante d'Adobe Acrobat.

¹³ C'est-à-dire écrit en langage naturel, par opposition au langage contrôlé utilisé dans certains domaines ou entreprises pour limiter le vocabulaire spécifique et justement ainsi faciliter l'indexation.

construction de ces représentations expliquent le fait que ces structures sémantiques ne soient pas exploitées pour des grands corpus de documents. »

Nous ne contestons pas le fait que la mise en oeuvre d'un projet de gestion des connaissances (voir chapitre 3) représente un coût considérable pour une organisation. De plus, quel que soit le type d'outil opérationnel auquel ce projet permet d'aboutir (système à base de connaissances, base de connaissances ou livre de connaissances hypermédia), ceux-ci sont généralement aussi complexes dans leur fonctionnement que dans leur utilisation. Il serait donc absurde de vouloir initier la conception d'un tel outil dans le seul but de générer des index documentaires.

Cependant, le fait est que des projets de gestion des connaissances sont menés dans des organisations et aboutissent à de tels systèmes. Il existe par ailleurs de nombreuses activités qui permettent de décrire un domaine, par exemple sous forme d'ontologie, de thésaurus des termes spécifiques ou encore d'annotation des documents. Rien n'empêche alors d'utiliser ces outils pour générer l'indexation d'un corpus documentaire spécifiquement au domaine auquel il se rapporte. Une telle approche permet a priori d'obtenir des index ciblés sur un domaine, et donc conçus spécifiquement en fonction des besoins en information des individus y exerçant leur activité.

2.2.2.1.2. *Vers une approche conceptuelle de l'indexation*

L'indexation obtenue est alors fortement reliée au domaine. La diversité et la complexité de ces relations dépassent la simple attribution de rôles. On parle alors d'*indexation conceptuelle* du corpus documentaire. Ce type d'indexation, s'il est actuellement de plus en plus utilisé, n'est que très rarement défini dans la littérature. À ce jour, nous n'avons trouvé qu'une seule définition, donnée par Yannick Prié ([Prié 01]) : « *une indexation offrant dans son mode de fonctionnement/représentation même la possibilité d'interroger et de manipuler la connaissance elle-même, en tant qu'indexation et plus seulement en tant qu'index pointant vers une information "brute"*. »¹⁴ À notre sens, la notion d'indexation conceptuelle est à rapprocher de celle d'identification de catégories (ou « document categorization », cf. [Massand & al. 92], via [Iwayama & Tokunaga 95], [Hearst & Karadi 97]) de documents ou

¹⁴ Il apparaît dans le reste de l'article de Y. Prié que le concept de « connaissance » utilisé n'est pas très clairement défini. Cet auteur dresse cependant un éventail des « ensembles de connaissances de description » qui peuvent permettre de construire des index conceptuels : « elles seront *a minima* de simples vocabulaires, et prendront *a maxima* la forme d'ontologies, c'est-à-dire de définition des concepts du domaine visé et de leurs relations, ainsi que de règles exprimant les possibilités de description. »

de celle de modélisation de l'information (ou « information modeling », cf. [Lucarella & Zanzi 96b], [Médini & al. 01]).

Les travaux auxquels Yannick Prié fait référence en parlant explicitement d'indexation conceptuelle sont récents et utilisent par exemple des systèmes d'annotation conceptuelle ([Zarri & al. 99]) ou des vocabulaires. Ceux-ci sont présentés sous forme de taxonomies (i.e. des thesaurus, [Erdmann & Studer 00]) ou d'ontologies génériques ([Guarino & al. 99], [Martin & Eklund 00]) ou spécifiques aux domaines ([Motta & al. 00]).

Il existe cependant des travaux plus anciens relevant à notre sens également de l'indexation conceptuelle. C'est par exemple le cas des travaux sur l'extraction automatique de descripteurs (ou *marqueurs*) linguistiques pour la structuration de corpus, à partir de dictionnaires ([Ahlsvede 88], [Vossen & al. 89]) ou par exploration contextuelle, comme le système SEEK (Système Expert d'Exploration Contextuelle, [Jouis, 1993]) qui a pour but de trouver des relations statiques entre les objets du domaine des comptes-rendus d'expertise. Les travaux de Jean-Luc Simoni se fondent sur l'extraction de terminologies pour la construction automatique de graphes de termes pour la recherche documentaire ([Simoni 00]). [Assadi & Gros 99] présentent eux des « Systèmes de Consultation de la Documentation Technique (SCDT) qui s'appuient sur des modèles conceptuels de structure arborescente (un modèle du domaine, proche de celui des concepts MKSM et un modèle des tâches, qui est en fait une ontologie des différentes tâches répertoriées dans l'activité modélisée). Ces SCDT sont d'abord construits automatiquement par des outils de traitement automatique de la langue directement appliqués à la documentation à indexer, puis affinés et validés par un cogniticien et un expert.

Parmi ces travaux, [Martin & al. 00], [Motta & al. 00] et surtout [Assadi & Gros 99] sont les plus proches de notre problématique. Cependant, ils ne s'appuient que sur des représentations graphiques arborescentes de type graphe conceptuel. Il est à noter qu'à notre connaissance, aucun travail ne fait référence à des structures diagrammatiques plus complexes, comme on en rencontre dans la méthode MKSM, et mettant en jeu des relations sémantiques autres que des relations de spécialisation ou d'attribut.

Dans son travail sur les annotations¹⁵ de documents, [Prié 01] aborde cependant ce problème : il s'agit pour cet auteur de « rajouter du contenu » aux documents par des mécanismes d'annotation complexes et parfois non linéaires (une annotation peut faire référence à

plusieurs parties d'un document ou à une autre annotation). Techniquement, il résout ce problème grâce à la possibilité, en XML, de définir des liens intra-documentaires (l'annotation est en effet partie intégrante du document). Toutefois, relier un document à une annotation – aussi complexe soit-elle – de ce même document nous semble toujours plus aisé à réaliser que le lien entre un corpus documentaire entier et une collection de diagrammes structurée traitant du domaine d'application de ce corpus.

Le développement récent du nombre de travaux dans le domaine de l'indexation conceptuelle de documents témoigne de l'enjeu réel de la structuration « orientée-connaissances » des fonds documentaires dans les organisations. Comme dans tous les cas où l'on crée un nouveau concept, on a d'ailleurs regroupé par dichotomie tous les concepts existants dans un vocable opposé¹⁶. On parle désormais d'*indexation structurelle* pour qualifier tous les autres types d'indexation fondés sur la localisation des termes dans les documents d'un corpus et présentés au paragraphe précédent.

Cependant, [Prié 01] tend à atténuer cette distinction : pour cet auteur, toute structure documentaire contient des éléments de sémantique. Par là-même, elle relève du concept de connaissance et donc de l'indexation conceptuelle. Nous considérons toutefois dans ce qui suit que ce concept de connaissance, tel qu'il est envisagé en gestion des connaissances, est plus fortement relié au domaine d'activité de l'organisation qu'à la structure des documents qu'elle emploie. Dans le troisième chapitre de sa thèse, J-L. Simoni ([Simoni 00]), qui s'intéresse à l'étude des différents types de relations permettant de structurer un corpus, classe également ces relations en deux catégories distinctes : structurelles et sémantiques.

D'un point de vue fonctionnel, les représentations obtenues par ces deux types d'indexation n'ont pas la même opérationnalité. Dès lors, et même si indexation structurelle et conceptuelle sont de même niveau, nous continuons dans la suite à faire la distinction entre ces deux concepts.

2.2.2.1.3. *L'indexation de la requête*

L'utilisateur d'un moteur de recherche formule (ou formalise) une requête textuelle en utilisant un « langage de requêtes » dont la syntaxe lui permet d'associer plusieurs termes ou

¹⁵ L'auteur met en avant le fait que l'annotation est la forme la plus simple d'expression de la connaissance à propos d'un document. Dès lors, elle peut tenir lieu d'indexation conceptuelle de celui-ci.

¹⁶ Comme on a parlé de montre « à aiguilles » après l'invention de la montre à cristaux liquides, de courrier « postal » depuis l'explosion du courrier « électronique » ou de téléphone « filaire » par opposition au téléphone portable.

expressions. Il existe plusieurs types de tels langages. Les plus connus sont les langages booléens, qui permettent d'associer les différents termes ou expressions grâce à des opérateurs logiques ou le langage naturel, qui permet à l'utilisateur d'exprimer librement sa question. Entre ces deux extrêmes se situent des langages formels, comme le SQL, notamment utilisé pour l'interrogation de bases de données.

Une fois le corpus indexé, le système d'accès à l'information doit également réaliser l'indexation des requêtes. Mais si l'indexation d'un corpus se fait a priori, l'indexation de la requête doit elle être réalisée « à la volée ». Pour pouvoir réaliser un appariement de la question et des documents, les systèmes de recherche d'information doivent être capables d'indexer la requête d'une façon aussi proche que possible de celle du fonds documentaire.

Dans les cas simples, tels que l'indexation à plat, ce mécanisme consiste à générer dynamiquement la représentation de la requête attendue (i.e. une liste des termes de la requête pour l'indexation à plat). Dans le cas d'autres types d'indexation – et notamment pour l'indexation conceptuelle – les mécanismes employés sont parfois lourds (cf. paragraphe précédent). Ils peuvent nécessiter soit des temps de calcul importants, soit l'intervention d'un opérateur humain (comme c'est le cas avec la méthode MKSM) qui rendent irréalisable une indexation dynamique des requêtes.

D'une part, il est possible de demander à l'utilisateur de générer directement des requêtes compatibles avec le mode d'indexation utilisé. C'est ce que font Guarino & al. [Guarino 00], en lui faisant décrire des « graphes conceptuels lexicaux » qui sont appariés au graphe représentant le corpus documentaire, ou Martin et Eklund [Martin & Eklund 00] qui génèrent des « graphes-requêtes ». Cette solution oblige l'utilisateur à adhérer complètement à la représentation du corpus qui lui est présentée pour pouvoir effectuer sa recherche d'information. Ce pré-requis peut être vu comme un avantage ou un inconvénient en fonction du degré de perception du problème d'information. En effet, une telle représentation, si elle permet à un utilisateur ne connaissant pas le domaine (ou le connaissant suffisamment peu) de formuler une requête cohérente à partir d'éléments de la représentation proposée, elle peut en revanche entrer en conflit avec le modèle d'un utilisateur qui connaît déjà le domaine. Dans ce dernier cas, l'utilisateur peut être obligé de « désapprendre » pour pouvoir se reconstruire un modèle cohérent avec l'indexation présentée.

Une autre solution est de permettre à l'utilisateur de formuler une requête sous forme textuelle. Le système doit ensuite réaliser l'appariement de cette requête avec l'indexation conceptuelle du corpus. Erdmann et Studer ([Erdmann & Studer 00]) réalisent cela grâce à un

système de règles et à un moteur d'inférences. L'intérêt de cette solution est qu'une fois mis en place, ce système permet de court-circuiter la phase d'indexation de la requête, puisque celle-ci n'a plus à être appariée à l'index conceptuel car la requête est directement projetée sur celui-ci¹⁷. D'où un gain de temps évident dans les réponses du système. Cependant, l'inconvénient d'un tel système est qu'en surface (i.e. pour l'utilisateur), il n'est pas différent d'un SRI classique. En effet, l'utilisateur n'a pas accès à l'indexation conceptuelle réalisée, et ne peut donc pas s'en servir pour formaliser son problème d'information.

Chacune de ces fonctions d'indexation documentaire aboutit donc à des représentations du fonds documentaire et des requêtes différentes. C'est pourquoi elles sont souvent utilisées de paire avec un type de fonction de comparaison dans le processus de recherche d'information. Le paragraphe suivant présente les différentes approches de conception de ce dernier type de fonctions.

2.2.2.2. La comparaison documentaire

Une fonction de comparaison a pour but d'apparier la représentation de la requête avec celle de chacun des documents du corpus. Techniquement, d'après [Roussey & al. 1999], il existe deux types de fonctions de comparaison. Les unes fournissent une liste de documents pertinents (avec éventuellement un tri) par rapport à la question ; ce sont les plus couramment utilisées – notamment dans les moteurs de recherche. Les autres trient les documents du corpus par ordre de pertinence. Dans le premier cas, le but est d'obtenir une sous-partie du corpus contenant les documents pertinents, tandis que dans le second, il s'agit de trier l'ensemble des documents du corpus.

Ces deux types de fonctions se distinguent également par des modèles de recherche d'information différents. Il existe plusieurs types de modèles de recherche d'information (booléen, vectoriel, probabiliste...). Ces modèles sont naturellement définis en cohérence avec la méthode d'indexation choisie. Par exemple, un modèle booléen correspond bien à une indexation à plat, car il s'agit de retrouver des chaînes de caractères et d'identifier la conjonction de celles-ci dans les documents du corpus. Les modèles vectoriels et probabilistes correspondent eux à une représentation du corpus par indexation pondérée.

Il nous paraît utile de faire également figurer dans ce paragraphe sur la comparaison documentaire des méthodes moins usuelles dans le domaine de la recherche d'information,

¹⁷ En fait, si l'indexation de la requête n'est pas réalisée formellement par le système, elle est sous-tendue par le système de règles qui a été élaboré en fonction d'un modèle de l'ensemble des requêtes documentaires possibles.

mais qui ont été utilisées dans les systèmes présentés au paragraphe précédent. En effet, les mécanismes d'indexation conceptuelle s'appuient parfois sur des représentations particulières et il convient alors de mettre au point une méthode de comparaison spécifique à ces systèmes.

En particulier, la plupart des travaux fondés sur des structurations arborescentes du domaine réalisent directement l'appariement des graphes obtenus pour représenter le corpus et les requêtes en utilisant des techniques de projection géométrique ([Martin & al. 00]), de subsomption ([Guarino 00]) ou d'extension des requêtes ([Zarri & al. 99]). Pour les représentations plus complexes (i.e. non arborescentes) ou les cas où les représentations du corpus et des requêtes ne sont pas réalisées de la même manière, la technique généralement utilisée est l'emploi de règles d'inférences ([Erdmann & al. 00], [Heflin & al. 99], [Jouis, 1993]), qui permettent de définir un éventail de relations plus vastes que la correspondance directe recherchée dans l'appariement. En présentant une plus grande adaptabilité, ces règles permettent à la fois de mettre en correspondance des représentations structurellement et sémantiquement éloignées et de traiter les « cas particuliers » (expressions idiomatiques des requêtes, structures récursives des graphes...).

2.2.2.3. Les critères d'évaluation

Les différentes techniques d'indexation, d'interrogation et de comparaison conduisent à des différences de performance entre les systèmes. Il s'agit ici d'indiquer les méthodes permettant l'évaluation de ces performances. Il est à noter que ces méthodes ne sont pas spécifiques aux systèmes à base d'interrogation, et peuvent être utilisées quel que soit le type de système employé.

Ces méthodes s'appuient sur un corpus de test, dont le nombre total de documents (N) et les documents pertinents par rapport à une requête (dont le nombre est Np) sont connus. Le nombre de résultats (i.e. de documents retournés par le système) est noté Nr. Parmi ces résultats, le nombre de documents pertinents est noté Nrp. La performance d'un système de recherche d'information peut être mesurée en termes de bruit/silence ou de rappel/précision.

Dans le premier cas, l'idée est de déterminer les imperfections du système. Le bruit et le silence mesurent respectivement les proportions de documents non pertinents retournés et pertinents non retournés par le système par rapport à la taille du corpus. Du point de vue de l'utilisateur, le bruit représente alors « ce qui gêne » et le silence « ce qui manque ». Ces paramètres s'écrivent :

$$\text{Bruit} = (Nr - Nrp) / N \quad \text{et} \quad \text{Silence} = (Np - Nrp) / N$$

Le rappel et la précision donnent la mesure de la performance du système (i.e. le nombre de documents pertinents retournés) en fonction respectivement du nombre de documents théoriques à retourner et du nombre de réponses du système.

$$\text{Rappel} = \text{Nrp} / \text{Np} \quad \text{et} \quad \text{Précision} = \text{Nrp} / \text{Nr}$$

En règle générale, les critères de rappel et de précision, même s'ils sont conceptuellement moins évidents à définir, sont plus utilisés pour l'évaluation des systèmes de recherche d'information, car ils permettent d'accéder directement à une mesure des performances du système.

2.2.3. Les stratégies de parcours

Une alternative aux stratégies analytiques d'accès à l'information est proposée par les stratégies de parcours, encore appelées *browsing* ou *navigation*. Pour Vannevar Bush, ces stratégies correspondent « à la nature associative de l'esprit humain » ([Bush 45], via [Balpe & al. 96], p. 11). Ces stratégies s'appuient sur la démarche de tâtonnement et d'approche de l'information recherchée de proche en proche. L'emploi le plus courant de ce type de stratégies est la navigation hypermédia (voir paragraphe suivant).

Cette idée de tâtonnement est à rapprocher de celle de « butinage » (ou « berrypicking » de [Bates 89], illustré en Figure 5). Tel un individu qui recherche de la nourriture, celui qui cherche de l'information en utilisant une stratégie de parcours va se diriger vers les endroits qui lui semblent les plus intéressants pour sa recherche et lorsque chacun d'eux est épuisé, il en repère un autre à partir de l'endroit où il se trouve. Pour cela, il ne suit pas une trajectoire préétablie et peut repasser plusieurs fois au même endroit.

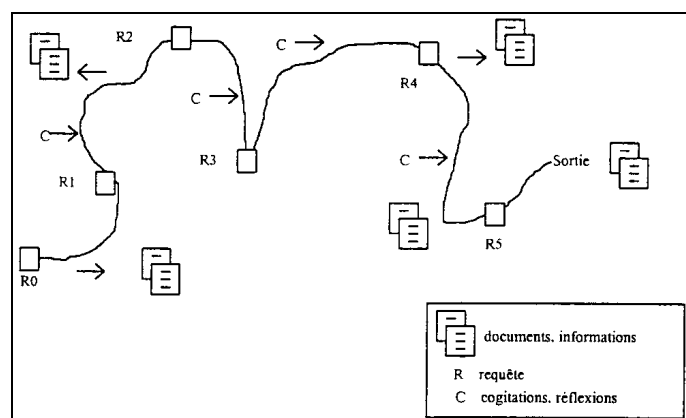


Figure 5. Représentation de l'accès à l'information en « cueillette de baies » de M. Bates.

Cette analogie présente toutefois ses limites. En effet, le but d'une stratégie de parcours (i.e. la solution du problème d'information) réside dans la dernière étape de ce parcours. Plus précisément, la recherche d'information se termine dès que l'individu qui emploie cette stratégie estime pouvoir répondre par l'affirmative à la question fermée « Cette étape de la navigation est-elle solution de mon problème ? »¹⁸ Celui qui réalise une tâche de cueillette s'intéresse lui à chaque étape de son cheminement, qui est censée apporter un élément de réponse quantitatif à la solution globale de son « problème d'alimentation ».

Le terme générique de cheminement emprunté – qui correspond à la navigation dans le contexte de l'accès à l'information – désigne la somme des déplacements réalisés pour accéder à chaque étape du parcours. Chacun de ces déplacements est le résultat d'un choix fait par l'individu de passer d'une étape à une autre de son cheminement. Ce choix est influencé par la perception courante du but de son cheminement (ou de son problème d'information), dont nous avons vu qu'elle est modifiée par le résultat des étapes précédentes, et de la vision qu'a l'individu des étapes potentielles suivantes. Pour Gary Marchionini, ces stratégies de parcours sont à la fois heuristiques¹⁹ et opportunistes ([Marchionini 95], p. 8).

Plusieurs types de stratégies de parcours sont distingués dans la littérature. Les stratégies linéaires ou de butinage de [Claverie & al. 97] ainsi que les stratégies spatiales et conceptuelles de [Foucault 99] sont des stratégies heuristiques génériques qui peuvent être employées indépendamment de l'espace dans lequel se déplace l'individu. D'autres types de stratégies non optimales peuvent être définis en prenant en compte les spécificités de cet espace. Ces spécificités peuvent être représentées par un typage particulier des étapes potentielles de cet espace ou des liens les reliant. Lors d'un déplacement dans un réseau routier, décider de faire une étape dans une ville particulière ou choisir d'éviter les cols d'altitude ou d'emprunter l'autoroute sont des exemples de telles stratégies. Dans tous les cas, ces stratégies sont identifiables au niveau global du cheminement parcouru dans l'espace considéré. [Claverie 97] parle donc de « styles de navigation » pour caractériser ces stratégies.

¹⁸ Nous supposons ici que l'information recherchée existe, est accessible et est unitaire (i.e. elle n'est pas disséminée dans plusieurs éléments de la structure informationnelle présentée).

¹⁹ Selon [Sakarovitch 84], « Un algorithme est dit "approximatif" ou "heuristique" s'il conduit toujours à une solution réalisable, mais pas nécessairement à une solution optimale. » Cette définition, issue de la recherche opérationnelle, s'applique d'après son auteur également à « l'Intelligence artificielle en général ». Dans ce document, nous élargissons cette définition d'une heuristique à celle d'une stratégie de résolution d'un problème non nécessairement optimale, et ne menant pas nécessairement à la solution de ce problème. Nous considérons en effet qu'un individu peut employer une stratégie heuristique d'accès à l'information sans trouver l'information cherchée.

L'identification de ces styles se fait à l'aide d'une carte de l'espace parcouru. Nous revenons au paragraphe 2.2.4.1 sur l'intérêt des cartes pour l'élaboration de telles stratégies.

Il nous semble par ailleurs qu'il peut exister des stratégies de parcours non heuristiques, correspondant à un cheminement optimal dans l'espace pour la résolution du problème. C'est notamment le cas lorsque l'intérêt d'un cheminement par rapport aux autres est évident. L'exemple le plus évident est la ligne droite, reconnue comme étant le plus court chemin entre deux points.

Dans tous les cas, cette optimalité (et plus généralement la pertinence de la stratégie adoptée) est définie en fonction d'un ou de plusieurs paramètres de « mesure de l'espace parcouru »²⁰. Le temps nécessaire au parcours, la « distance » totale parcourue, le nombre d'étapes ou l'intérêt de ces étapes peuvent par exemple être quantifiés pour permettre cette mesure. L'évaluation de la pertinence d'une stratégie heuristique de parcours nécessite alors l'identification des paramètres de mesure pertinents.

2.2.4. *Systèmes hypertextes et hypermédia*

L'idée d'un système informatique capable de reproduire une démarche de cheminement dans une structure d'information non linéaire est attribuée à Vannevar Bush avec son système Memex ([Bush 45]). Ce système devait permettre d'explorer et d'annoter de l'information textuelle ou graphique. Ce n'est qu'une vingtaine d'années plus tard que Douglas Engelbart met au point le système NLS (oN Line System), qui sera la première implémentation de l'idée de Bush ([Engelbart 62]). Quelques années plus tard, Theodor Nelson invente les termes d'hypertexte et d'hypermédia ([Nelson 67]) et imagine le projet Xanadu. Pour cet auteur, ce système devait permettre d'améliorer les capacités intellectuelles de son utilisateur en matière de manipulation d'informations et de produire des inférences sémantiques permettant la réutilisation des informations existantes dans la production de nouvelles informations (il parle de « transclusion » ou de « transédiction », cf. [Alix & Roussel 01]). Cette idée devait déboucher « à la limite, sur l'utopie d'une "bibliothèque universelle". » Mais cette conception des hypermédias, parfois qualifiée de « romantique » [Tricot & Nanard 98], a depuis lors été supplantée par une vision plus opérationnelle des possibilités de ces systèmes, notamment en termes de présentation de l'information. Les développements des interfaces graphiques sur les postes de travail individuels et la diffusion des premiers systèmes hypermédias « grand

²⁰ Au sens de la théorie des mesures en géométrie.

public »²¹ ont largement contribué à l'utilisation de l'hypermédia comme mode usuel de présentation de l'information. La standardisation des techniques de présentation est notamment due à la généralisation de l'utilisation de l'Internet.

Des communautés scientifiques se sont créées autour des différents aspects des hypermédias, comme le World Wide Web, les environnements de conception d'hypermédias, les systèmes hypermédias d'accès à l'information (Hypermedia Information Retrieval Systems, ou HIRS, les applications de l'hypermédia à l'apprentissage ou les hypermédias adaptatifs (encore appelés Documents Virtuels Personnalisables [DVP]). La littérature abonde dans chacun de ces domaines, autant au format papier que sur le web²².

La navigation hypermédia est donc devenue une activité usuelle permettant une exploration non linéaire des données mises à disposition²³. Dans la perspective de l'accès à l'information, les systèmes hypermédias permettent une interaction directe avec ces données. Cette interaction a toutefois mis au jour un problème particulier lié à la désorientation des individus dans la structure des données.

Ce paragraphe s'intéresse aux différents aspects de la conception de systèmes hypermédias d'accès à l'information. Une description générique des systèmes hypertextes est tout d'abord présentée. Puis, nous évoquons le problème de désorientation spécifique à ces outils et l'identification des solutions envisagées pour le résoudre. Les aspects propres à la conception de systèmes hypermédias d'accès à l'information sont ensuite présentés.

2.2.4.1. Modèles d'hypertextes

Un hypertexte et un hypermédia sont bâtis sur le même modèle. Ils se composent de trois éléments.

Les nœuds représentent le contenu informationnel. Ils sont constitués d'unités textuelles (pour un hypertexte) ou multimédias (pour un hypermédia)²⁴. Chaque nœud contient un

²¹ Le premier d'entre eux étant le système Hypercard®, diffusé par Apple™ en 1987.

²² Il est par ailleurs important de noter que si les écrits papiers sont soumis à la validation de la communauté intéressée, la qualité des références que l'on peut trouver sur l'Internet – hormis sur les sites officiels de ces communautés – est très variable.

²³ [Golovchinsky 97] (p. 407) classe d'ailleurs cette activité « into the lay public's almost implicit consciousness », ce qui ne nous semble pas très éloigné de la notion d'« inconscient collectif ».

²⁴ Dans la suite de ce paragraphe, nous utilisons tout à tour les termes d'hypertexte ou d'hypermédia, pour alléger le style et adopter la terminologie des auteurs auxquels nous faisons référence. Les considérations émises sont également valables pour ces deux types d'applications, sauf indication contraire. Le terme de « visualisation » est à considérer, pour de l'information multimédia, comme une analogie avec son utilisation pour l'information textuelle. Il s'agit d'y accéder via un outil approprié.

élément d'information supposé unitaire, c'est-à-dire qu'il permet de véhiculer une « idée » dans sa totalité, qui peut toutefois se définir en fonction d'autres nœuds.

Les liens permettent le passage d'un nœud à un autre lors de la navigation. Ils sont associés à une relation sémantique entre les nœuds, entre des éléments d'un nœud ou entre des éléments de deux nœuds. Les liens peuvent être unidirectionnels ou bidirectionnels. Dans cette thèse, seuls les liens unidirectionnels sont pris en compte. Les liens bidirectionnels sont alors considérés comme deux liens unidirectionnels matérialisant deux relations sémantiques réciproques. Un lien unidirectionnel permet le passage d'un nœud source à un nœud cible de l'hypertexte.

Les ancres sont les éléments des nœuds à partir desquels sont accessibles les liens. Le fait qu'une ancre soit un sous-élément d'un nœud permet de définir un lien comme une relation relative à une « unité sémantique de niveau inférieur à celui du nœud » ([Balpe & al. 96], p. 21).

En pratique, le stockage des nœuds d'un hypertexte se fait en général²⁵ dans des fichiers séparés. Dans le cas d'un hypermédia, un nœud et un élément multimédia de ce nœud peuvent également être stockés séparément. Les liens sont définis en tant que références à partir des ancres, c'est-à-dire à partir d'éléments des nœuds. Un « fichier-nœud » contient donc des éléments textuels ou des références aux éléments multimédias, l'identification des éléments considérés comme des ancres et les références aux cibles des liens correspondants. Les deux types de références considérées ici peuvent être formalisées de la même façon (voire concerner les mêmes objets), mais sont traitées différemment : les références multimédias aux éléments du nœud sont visualisées dans celui-ci, alors que celles situées dans les liens ne sont accédées que lors du parcours (on parle également d'activation) du lien correspondant.

Plusieurs approches formelles ont été adoptées pour permettre de structurer les hypertextes (voir paragraphe suivant pour les raisons qui président à cette structuration). [Nanard & Nanard. 98] identifient notamment cinq dimensions de l'espace de conception d'hypermédias : le produit, le processus de conception, le domaine, l'équipe de conception et l'utilisateur. Le domaine est constitué des documents (i.e. l'information) et des connaissances (au sens de [Nanard & Nanard 91], p. 330, c'est-à-dire des connaissances procédurales²⁶ directement applicables à la structure de l'hypertexte). Cet aspect, ainsi que celui de l'équipe

²⁵ C'est-à-dire dans le cas des formats d'hypertextes dédiés au World Wide Web, qui nous intéressent dans cette thèse. Les aspects techniques de ces formats sont présentés au chapitre 3.

²⁶ Cf. [Ermine & al. 96].

de conception sont des composantes de l'approche de cette thèse. Ils sont décrits dans les perspectives méthodologiques au chapitre 3. La prise en compte de l'utilisateur est discutée dans les perspectives ergonomiques de ce chapitre, au paragraphe 2.3.4.

La définition d'un processus de conception spécifique à l'approche retenue est l'un des objectifs de ce travail de thèse. Elle n'est pas discutée ici dans le contexte des hypermédias, mais élaborée dans le contexte plus global de l'accès à l'information par recontextualisation. Signalons toutefois la démarche générale donnée par ces auteurs, qui s'appuie sur un modèle abstrait de l'application et des instances de celles-ci. Le prototypage est un élément important de l'activité « opportuniste » de conception de ces instances, via les cheminements ascendant et descendant de conceptualisation et d'instanciation. Nous prenons également note de la nécessité pour ces auteurs de mener une conception « raisonnée » (au sens du Design Rationnel, voir chapitre 3) de l'application hypermédia, en justifiant les choix de conception. Nous discutons cependant plus loin la pertinence de l'argumentation graphique dans le processus de conception.

Concernant la perspective produit, ces auteurs (p. 21) recensent plusieurs types de modèles. Parmi ceux-ci, les modèles génériques, tels que le modèle Dexter ([Halasz & Schwartz 94]) ou son extension, le modèle Amsterdam ([Hardmann & al. 94]) permettent de spécifier des systèmes indépendamment de leurs contenus. Leur application est notamment utile pour les systèmes dont la structure est destinée à être réutilisée. D'après ces auteurs, les modèles de conception, comme HDM ([Garzotto & al. 91], [Garzotto & al. 93]), apportent « une perspective méthodologique » pour la conception d'hypertextes et d'applications hypermédias. Les modèles de l'hypermédia cible et d'implémentation sont destinés à spécifier l'hypertexte à réaliser en termes de fonctionnalités et de caractéristiques techniques.

Jean-Pierre Balpe et alii évoquent également d'autres formalismes, qu'ils qualifient de « modèles dynamiques », au sens où ils permettent notamment de spécifier le comportement temporel et la synchronisation des hypertextes (ouvertures et fermetures de fenêtres, apparition et disparition d'informations). Ces approches sont fondées notamment sur les automates à états finis, les réseaux de Pétri ou les hypergraphes ([Balpe & al. 96], p. 41). Ces auteurs notent cependant que ces formalismes ne s'appliquent qu'à des contextes particuliers, parmi lesquels les applications documentaires usuelles ne figurent pas.

Il nous semble toutefois que l'opportunité d'un modèle fonctionnel, soulignée plus haut, coïncide avec cette notion de modèle dynamique. Indépendamment des formalismes adoptés,

il est alors nécessaire de disposer d'une vision structurelle et d'une vision fonctionnelle de l'hypertexte à élaborer.

Concernant les aspects dynamiques des applications hypermédia, différentes approches de réalisation existent. Nous ne retenons ici que celle des Documents Virtuels Personnalisables (DVP, cf. [Crampes 99]), dont les aspects techniques sont décrits au chapitre 3. Cette approche, qui a émergé en France lors de la conférence IHM'99, est à rapprocher du courant des « Adaptive Hypermédia » ([Brusilovsky 96]). Toutefois, cette approche (et la technologie sous-jacente) est issue d'autres travaux sur la réutilisation des structures de données, comme ceux de [Nanard & al. 98].

Dans cette thèse, nous ne faisons pas de différence entre liens statiques et liens dynamiques, comme cela est parfois fait dans la littérature. Nous associons cette caractéristique aux nœuds plutôt qu'aux liens. Les liens dynamiques sont alors considérés comme les liens statiques d'un hypertexte généré dynamiquement.

2.2.4.2. Désorientation et cartes conceptuelles

Les stratégies de parcours induisent une *surcharge cognitive* qui peut être vue à deux niveaux. À chaque étape de la navigation, le browsing requiert pour l'individu le choix de sélection d'un lien. Au niveau navigationnel, il se construit en même temps une représentation de la structure des données, de sa position et de sa navigation passée et future dans cette structure²⁷ ([Nielsen 90], p. 127). Cela est valable quel que soit le type de navigation utilisé par cet individu. Lorsque cette surcharge est trop importante, elle aboutit au syndrome de désorientation (« *lost in hyperspace* ») introduit par [Conklin 87] et désormais couramment admis.

Pour [Balpe & al. 96] (p. 31), ce problème est entre autres lié à la taille de la structure de données représentée. Indépendamment de ce paramètre, [Nanard & Nanard 91] se réfèrent à [Bernstein 90] pour affirmer qu'il est lié à un manque d'organisation de la structure de

²⁷ Pour illustrer cette surcharge cognitive à travers les différentes stratégies de navigation, nous prenons l'exemple de la lecture d'une note de bas de page suffisamment volumineuse comme celle-ci. Du fait qu'elle ne s'insère pas dans la continuité du texte, elle induit une rupture de séquence et donc un parcours physique du centre au bas de la page. Cette note peut être consultée immédiatement après la lecture de l'appel de note. Cela oblige le lecteur à revenir au début de la phrase dans laquelle se situe l'appel après avoir fini de lire la note. Il peut également décider d'attendre la fin de la phrase, voire du paragraphe pour effectuer le saut. Dans ce cas, il doit mémoriser, jusqu'à la fin de sa lecture, le fait qu'il a un parcours de lien vers une note de bas de page « en attente ». Dans les deux cas, il doit focaliser en même temps son attention sur l'information signifiante du texte et sur l'information de parcours.

l'hypermédia. Il s'agit en fait de permettre la compréhension par le lecteur de l'organisation de cette structure.

Dans d'autres travaux, ces auteurs proposent de résoudre ce problème en identifiant les « connaissances factuelles » liées aux ancrs de liens ([Nanard & Nanard 93]). Le typage sémantique de ces ancrs (ainsi que celui des liens) et la prise en compte de leur contexte doivent permettre une structuration porteuse de sens pour le lecteur et une navigation au niveau des connaissances plutôt qu'à un niveau documentaire. L'idée sous-tendue est la composition personnalisée de l'hypertexte par rapport aux besoins de l'utilisateur, qui préfigure l'approche des Documents Virtuels Personnalisables décrite au chapitre 3.

Toutefois, une structuration complexe ou une taille trop importante de l'hypermédia sont parfois nécessaires. Une technique alors couramment employée pour pallier la désorientation est la présentation à l'utilisateur de la structure des données, notamment grâce à des *cartes conceptuelles* de l'hypermédia visualisé. De nombreux travaux ont été réalisés sur ces cartes, dont un inventaire est présenté dans [Roger & al. 98]. Il ressort de cette analyse que la technique de mise au point de cartes conceptuelles est souvent dérivée des recherches sur les méthodes de classification en sciences cognitives. La plupart de ces cartes proposent alors des taxinomies des domaines représentés (voir chapitre 3).

La présentation la plus courante de ces cartes est la mise à disposition dans une zone particulière de l'écran d'un ensemble de liens structurés de façon arborescente linéaire²⁸ et donnant accès aux différents *nœuds* de l'hypermédia²⁹. D'autres solutions sont cependant envisagées, comme la présentation des données en deux dimensions, comme [Novak & Gowin 84], ou la présentation interactive de ces cartes ([Zizi & Beaudoin-Lafon 94]).

La problématique de conception des hypermédiats est fortement influencée par la prise de conscience de ce problème de désorientation. L'activité de navigation est largement étudiée dans la littérature, sous l'angle des processus cognitifs mis en œuvre par l'interaction des utilisateurs avec ces systèmes. Ces travaux sont évoqués dans le paragraphe 2.3.4.1.

²⁸ La présentation des branches de l'arbre est « préordre », c'est-à-dire qu'un élément de niveau supérieur est présenté avant ses fils. Les différents éléments sont affichés en colonne. L'indentation permet d'identifier leur niveau de dans l'arborescence.

²⁹ Voir paragraphe suivant pour la définition des termes lien et *nœud*.

2.2.4.3. Les systèmes hypertextes d'accès à l'information

Un système hypertexte est a priori toujours destiné à l'accès à l'information, puisqu'il oblige l'utilisateur à prendre connaissance de cette information pour mener à bien sa navigation. [Tricot & Nanard 98] notent d'ailleurs le manque d'études spécifiques à ces systèmes. C'est pourquoi nous ne présentons pas d'état de l'art particulier dans ce paragraphe. Des systèmes permettant à la fois navigation et interrogation sont cependant étudiés plus loin.

Il nous semble en revanche nécessaire de définir ici ce que nous entendons par systèmes hypermédias d'accès à l'information, par opposition aux autres usages de ces systèmes. En effet, [Roger & al. 98] notent que la frontière entre hypertextes d'apprentissage et hypertexte d'accès à l'information est difficile à définir. Nous définissons dans le contexte de cette thèse un système hypertexte d'accès à l'information comme une application hypertexte destinée à permettre l'accès à l'information scientifique et technique à des utilisateurs dans le cadre de l'exercice de leur activité professionnelle. Cette définition, volontairement restreinte,

[Tricot & Nanard 98] étudient le processus de conception de ces outils sous l'angle de la modélisation des tâches des utilisateurs. Trois formalismes sont comparés et ces auteurs mettent en avant le fait que les stratégies décrites dans ces formalismes sont optimales, ce qui n'est pas toujours le cas des stratégies des utilisateurs. Ces auteurs notent également la variabilité de l'activité menant à la tâche de recherche d'information et constatent l'absence dans la littérature de travaux génériques prenant en compte la description de cette activité. Une approche ergonomique de l'accès à l'information par la description de l'activité contextuelle mérite donc d'être étudiée.

2.2.5. Un point de terminologie

Nous avons traduit plus haut la locution « *information retrieval* » par « accès à l'information ». Littéralement, cette locution désigne plutôt l'identification et la mise à disposition de l'information préalable à l'accès à cette information. Cependant, compte tenu du fait que cette activité n'a en général de sens que si elle est menée dans le but d'accéder à l'information³⁰, son acception la plus courante inclut l'activité d'accès à l'information (i.e. sa visualisation, pour une information textuelle). Inversement, l'activité d'accès à l'information

³⁰ Les seuls cas que nous ayons recensés où l'information retrieval se conçoit sans accès à l'information concernent l'évaluation de l'efficacité des outils destinés à cet accès. Cette évaluation peut être menée lors de la conception de ces outils ou de « campagnes d'évaluation » visant à comparer leurs performances respectives. La plus connue de ces campagnes est menée pour la conférence TREC (Text REtrieval Conference).

étant de peu d'intérêt en soi, nous incluons dans l'emploi de cette locution l'ensemble des processus préalables permettant cet accès, c'est-à-dire l'information retrieval.

Pour Gary Marchionini, *information seeking* correspond au processus global d'accès à l'information – incluant la définition du problème d'information. [Rouet & Tricot 98] ont traduit en français cette locution par « recherche d'information », en précisant qu'ils lui attribuent « un sens psychologique assez éloigné de celui que lui donnent habituellement les informaticiens ». Comme ce « sens psychologique » nous semble correspondre à la définition donnée ci-dessus, nous traduisons ici également *information seeking* par *accès à l'information*.

Gary Marchionini établit par ailleurs une distinction entre *information seeking* et *information search*. D'après cet auteur, le terme *search* désigne l'activité (« behavioral manifestation ») induite par ce processus. Il note d'ailleurs qu'une machine peut réaliser une activité d'*information search*, mais que l'*information seeking* est le fait d'opérateurs humains. Nous n'avons pas trouvé de traduction couramment employée dans la littérature pour « information search ». Toutefois, nous pouvons rapprocher cette activité générique de la notion de « tâche cognitive », telle qu'elle est définie dans [Tricot & Nanard 98]. Bien que ces deux notions ne se situent pas aux mêmes niveaux cognitifs, il nous semble qu'un parallèle peut être établi entre *information seeking* et *recherche d'information*. En effet, pour ces auteurs, la recherche d'information (appliquée au domaine des hypermédias) est une sous-tâche d'une tâche générique³¹, qui s'inscrit dans l'accomplissement d'une activité globale. La notion d'activité permet également une décomposition structurelle (au sens des formalismes MKSM, voir chapitre 3). Il nous semble alors possible de faire correspondre les niveaux de granularité de la notion de tâche de recherche d'information et de celle d'activité d'information seeking. Cette dernière désigne alors le contexte immédiat d'exécution d'une tâche de recherche d'information.

Le terme « querying » (littéralement : questionnement) correspond à l'action de formuler et de soumettre une requête à un individu ou à un système d'accès à l'information par interrogation. Nous le traduisons ici par *interrogation*.

Concernant les stratégies de parcours, nous nous référons à la distinction générique faite par Nigel Woodhead entre browsing et navigation ([Woodhead 90], p. 102). Pour cet auteur, le browsing représente le parcours au niveau microscopique, c'est-à-dire l'enchaînement des

liens et la modification du but pendant le parcours. C'est la stratégie de l'utilisateur qui choisit un lien à parcourir en arrivant sur un nouveau nœud. La navigation est à considérer à un niveau macroscopique : elle désigne l'ensemble du cheminement de l'individu dans la structure informationnelle parcourue, en fonction de son but de navigation. On parle donc de navigation pour qualifier un cheminement dans un hypertexte, tandis que l'on considère que l'utilisateur adopte une stratégie de browsing pour réaliser ce cheminement. Cette distinction, qui présente l'avantage d'être indépendante de la notion de but et plus généralement du contexte de l'accès à l'information, est utilisée dans la suite de ce manuscrit.

2.2.6. Spécificités des deux types de stratégies

Dans ce paragraphe, nous évaluons les spécificités de chacun des deux types de stratégies d'accès à l'information. Cela doit nous permettre d'évaluer la pertinence de chacun d'eux en fonction des différents types de démarches d'accès à l'information. Les critères de différenciation identifiés ci-dessous sont issus de [Marchionini 95] (p. 8).

Le premier de ces critères est le niveau de perception du problème d'information. Dans le cas de l'utilisation d'un système d'accès à l'information par interrogation, plus ce niveau de perception est élevé, plus clairement le problème d'information est exprimé et meilleures sont les réponses. À l'inverse, un individu ayant peu de connaissances dans un domaine peut poser des questions plus générales qu'un expert du domaine. Dans le premier cas, l'ensemble des réponses possibles à la question qu'il pose est alors plus vaste et celles-ci sont moins pertinentes pour son problème que s'il avait mieux réussi à délimiter le champ de sa recherche d'information. Dans le deuxième cas, il ne trouvera pas la réponse à sa question. L'intérêt d'une stratégie de browsing est ici manifeste : idéalement, une telle stratégie va le guider jusqu'à la solution de son problème sans qu'il ait à faire la démarche de formaliser celui-ci.

D'un point de vue cognitif, si une démarche interrogative requiert un effort préalable de formalisation du problème d'information, elle est ensuite moins « cognitivement » coûteuse qu'une stratégie de parcours³² (voir plus haut).

Par ailleurs, l'accès à une information précise en n'employant que des stratégies de parcours peut nécessiter un grand nombre d'étapes. C'est pourquoi il est communément admis que les stratégies analytiques sont plus rapides pour l'accès à l'information dans de « grandes

³¹ Ces auteurs citent les exemples des tâches d'information mining, manufacturing ou farming, qui correspondent aux activités d'utilisation de systèmes ayant des finalités différentes.

structures de données » ([Nielsen 90], p. 137, [Lesk 89], [Golovchinsky 97]). La taille critique à partir de laquelle un changement de stratégie s'impose est cependant floue. Personne n'imagine par exemple accéder à une information sur un serveur distant sur Internet à partir de la page de démarrage de son navigateur en cheminant de lien en lien sans passer par un moteur de recherche, même en supposant connues toutes les étapes de la navigation à suivre pour s'y rendre. Il n'est pas non plus courant de faire appel à un moteur de recherche pour atteindre une page web située sur un « petit » réseau local. Les deux types de stratégies sont par contre envisageables dans le cas d'une page située sur un « gros » réseau local ou sur un « petit » réseau rapidement accessible par navigation.

Une autre distinction faite par G. Marchionini est le niveau d'interactivité des deux types de stratégies : pour cet auteur, l'interrogation est plus formelle, dans la mesure où elle suit un schéma « *turn taking* » (« question/réponse »), alors que les stratégies de browsing sont « *more interactive, real-time exchanges, and collaborations between the information seeker and the information system.* » Nous nous permettons d'objecter à cela que, hormis dans le cas d'hypertextes dynamiques de type document virtuel personnalisable (DVP), l'information visualisée dans un hypertexte est figée³². La réponse d'un système à base d'interrogation est elle par nature dynamique et fonction de la question de l'utilisateur. Dans ces conditions, les niveaux d'interactivité de ces deux types de stratégies semblent pencher en faveur des stratégies d'interrogation. Il en va de même pour l'aspect « temps réel » de ces systèmes : la puissance des machines actuelles permet d'afficher plus rapidement les résultats d'une requête sur Internet que d'accéder à un site lointain. En revanche, l'activité de l'utilisateur d'un système à base de navigation est plus facile à aider ou à guider, notamment grâce à des « cartes conceptuelles » (voir plus loin) que celle de l'utilisateur d'un moteur de recherche.

Une autre distinction entre browsing et querying est le fait que seule cette dernière stratégie existe indépendamment de la recherche d'information qui la met en oeuvre. Une stratégie d'interrogation classique est par exemple une précision de la requête si le nombre de réponses est trop important ou sa généralisation s'il est trop faible. Pour ces raisons, G. Marchionini la définit comme planifiée (« *planned* »). À l'inverse, une stratégie de browsing est opportuniste et ne peut être définie qu'en fonction du contenu de l'information rencontrée à chaque étape de la navigation.

³² Pour G. Marchionini, « *browsing strategies demand a smaller cognitive load in advance and a steadier attentional load throughout the information seeking process.* ».

³³ Il convient ici de ne pas mélanger la nature statique ou dynamique de cette information, indépendante de l'outil d'accès à l'information utilisé, de son mode de présentation.

Cet auteur oppose également les caractères respectivement déterministe et heuristique du querying et du browsing. Pour avoir expérimenté le browsing déterministe aussi bien que le querying heuristique, nous nous permettons, une fois de plus, de tempérer cette dichotomie. Il nous semble que le caractère déterministe ou heuristique d'une démarche d'accès à l'information dépend plus du niveau de formalisation du problème d'information que de la stratégie employée. Il est toutefois exact qu'en permettant un accès à un plus grand nombre d'informations moins en rapport avec le but de la recherche, le browsing a plus tendance à induire des modifications du but de la recherche que le querying, pour lequel toutes les réponses présentées ont *en général* un rapport avec la question posée.

Il apparaît donc que plusieurs critères permettent d'orienter le choix d'un type de stratégie d'accès à l'information. Cependant, toujours pour Gary Marchionini, il n'y a pas de séparation nette entre les valeurs extrêmes de ces critères. Au contraire, chacun d'eux délimite un continuum aux extrémités duquel se situent querying et browsing. Ces continua sont illustrés dans le tableau ci-dessous, inspiré de la figure située en page 73 de [Marchionini 95].

Types de stratégies	Analytiques	Parcours
Niveau de perception du problème d'information	Formalisé	Conscient
Répartition de la charge cognitive	Avant la recherche	Pendant la navigation
Facteurs d'utilisabilité	Interactivité/ Rapidité	Aide Guidage
Taille de la structure de données	Grande	Réduite
Type de stratégie d'accès à l'information	Planifiée Déterministe	Opportuniste Heuristique

Tableau 1. Les continua de classification des stratégies d'accès à l'information.

Pour être exhaustif, nous devons noter que ces critères ne sont pas les seuls facteurs à prendre en compte pour déterminer le choix de la stratégie à adopter dans une démarche d'accès à l'information. Une telle étude est menée par [Roger & al. 98], qui s'appuient sur une étude bibliographique de neuf évaluations comparatives des stratégies d'interrogation et de parcours. Finalement, ces auteurs constatent qu'il est difficile de dégager un consensus sur les facteurs influençant réellement le succès des deux stratégies et mettent en évidence le manque d'un cadre fédérateur pour l'analyse de ces données. Ils concluent sur la nécessité d'analyser les différentes stratégies d'accès à l'information en termes de tâches cognitives (voir paragraphe 2.3.4.1). Une telle analyse (sur laquelle nous reviendrons) est présentée dans

[Tricot et Nanard 98], concernant les stratégies de parcours. Nous n'avons pas trouvé, d'étude de ce type concernant les stratégies analytiques, ni d'étude comparative de ces deux types de stratégies en termes de tâches.

Par conséquent, en l'absence d'un modèle précis de la littérature permettant l'identification précise des caractéristiques qui peuvent conduire un individu à choisir l'un des deux types de stratégies, nous choisissons l'exhaustivité de la définition de ces caractéristiques par Gary Marchionini : « *Information seeking depends on interaction between several factors: information seeker, task, search system, domain, settings and search outcomes* » ([Marchionini 89] et [Marchionini & Schneiderman 88], via [Marchionini 95], p. 32).

La conjonction de tous ces facteurs peut alors générer des situations où aucune des deux stratégies n'est totalement adaptée. Pour pallier cet inconvénient, une solution envisagée est de mettre en place d'autres stratégies d'accès à l'information, à partir d'une hybridation des stratégies d'interrogation et de parcours. Nous entendons ici par hybridation à la fois l'alternance des deux types de stratégies et la construction de nouveaux types de stratégies fondés sur les stratégies d'interrogation et de parcours.

2.2.7. Les stratégies alternées et les systèmes correspondants

Chacun des deux types de stratégies présente donc des avantages et des inconvénients. Étant donné les continua de situations d'accès à l'information auxquels les systèmes d'accès à l'information doivent répondre, il n'est pas évident de déterminer à l'avance quel type de stratégies est le plus à même de satisfaire l'utilisateur dans sa démarche d'accès à l'information.

D'après [Belkin & al. 93], il peut d'ailleurs exister un passage continu des individus d'une stratégie à l'autre. Une « synergie » entre les deux types de stratégies est alors cherchée par [Mackinlay & al. 95], et certains auteurs, comme Gene Golovchinsky, s'interrogent même quant à la pertinence de distinguer les deux types de stratégies ([Golovchinsky 97]). Ces auteurs s'appuient sur des systèmes spécifiques pour justifier ces points de vue. Nous restons cependant persuadé, à la lumière des distinctions faites dans la littérature par la majorité des autres auteurs, de l'intérêt de cette distinction ([Nielsen 90], chapitre 8, [Balpe & al. 96], chapitre 3, et plus généralement [Mc Knight & al. 91], [Marchionini 95], [Agosti & Smeaton 96]).

Toutefois, l'opportunité pour l'utilisateur de disposer des avantages des deux types de stratégies a été largement étudiée dans la littérature. La variabilité des différents paramètres

pris en compte dans ces études les rend à notre avis indissociables des solutions techniques mises en place pour les réaliser. Nous présentons donc dans ce paragraphe certaines de ces études, visant soit à laisser à l'utilisateur le choix de la stratégie à adopter, soit à explorer des stratégies d'accès à l'information alternant les deux types de stratégies, en parallèle avec les outils les mettant en œuvre.

Plusieurs exemples d'approches visant à laisser le choix aux utilisateurs de la stratégie d'accès à l'information à adopter sont référencés ici. Le système Hyperlaw présenté dans [Agosti & al. 91] s'appuie sur un modèle conceptuel à deux niveaux (nommé EXPLICIT) pour construire une représentation hypertextuelle de la base de données à interroger, qui est assez proche d'une représentation entité-association de la base de données. Le premier de ces deux niveaux est informationnel, et contient l'ensemble des documents de la base. Le second de ces deux niveaux permet de représenter les relations sémantiques entre les documents. Celui-ci peut être utilisé à la fois pour la navigation hypermédia dans la structure documentaire et pour l'interrogation du fonds. En ce sens, nous l'assimilons à un index conceptuel des documents, dans l'acception de ce terme définie plus haut.

De la même manière, [Salminen & al. 95] présentent un modèle formel de construction de systèmes hypertextes d'accès à l'information fondé sur un niveau informationnel et un niveau conceptuel. Ces auteurs identifient d'ailleurs clairement le niveau conceptuel comme un index des documents, et leurs travaux consistent en majeure partie sur la génération de cet index sous forme d'« hypergraphe ». Pour cela, ces auteurs s'appuient sur un découpage structurel des documents via une grammaire formelle appropriée. Le système généré permet alors à l'utilisateur une navigation spécifique (« expedite browsing ») dans l'hypergraphe ou une interrogation directe de l'index qu'il représente.

Le système DataStar Web ([Barker 97]) s'appuie sur la même stratégie d'accès à l'information. Nous n'avons pas d'informations précises sur le fonctionnement interne du deuxième système, mais il semble que celui-ci s'appuie sur deux représentations indépendantes des données (i.e. un modèle hypertextuel et une indexation à plat des données).

Ces trois systèmes permettent donc à l'utilisateur de choisir la stratégie la mieux adaptée selon lui à la résolution de son problème d'information. Toutefois, dans ces systèmes, le choix d'une stratégie est définitif. L'utilisateur désirant changer de stratégie en cours de recherche doit reprendre celle-ci du début. Ce type de système est donc pertinent si l'utilisateur est capable de choisir la meilleure stratégie à adopter pour sa recherche. Pour cela, il doit prendre en compte l'ensemble des paramètres évoqués au paragraphe précédent.

Une alternative à cette approche est de permettre à l'utilisateur de changer de stratégie en cours de recherche. C'est notamment l'approche adoptée par [Belkin & al. 93] pour le système BRAQUE. À l'instar des systèmes présentés plus haut, ce système s'appuie sur un modèle à deux niveaux permettant à la fois la recherche et la navigation par hypertextualisation dans les structures documentaires. Les auteurs mettent l'accent sur le passage continu d'une stratégie d'interrogation à une stratégie de parcours et sur l'aspect itératif de la démarche d'accès à l'information. Dans ce but, leur système décharge l'utilisateur de la tâche de formalisation de la requête en proposant une « progressive, interactive query formulation (and reformulation) ». Une des difficultés de cette démarche est d'imposer aux développeurs de concevoir deux systèmes pour la même finalité, qui ne doivent être ni trop éloignés ni trop proches. En effet, il s'agit d'une part d'assurer la cohérence pour l'utilisateur d'une stratégie d'accès à l'information alternée et de lui permettre cette alternance sans perte du bénéfice des étapes précédentes. D'autre part, des systèmes trop proches peuvent conduire l'utilisateur à s'interroger sur le choix de la stratégie à adopter, voire à être perturbé par une interface dont les composants font « double emploi ».

Une autre solution est l'emploi d'une stratégie fondée sur une séquentialisation des stratégies de browsing et de searching. Cette solution est notamment adoptée lorsqu'une stratégie se compose d'une phase de « restriction de l'espace de recherche » d'une phase d'« affinage » de la recherche (cf. [Balpe & al. 96], p. 95 ou [Lelu & al. 99]).

Par exemple, l'utilisateur d'un moteur de recherche sur le web peut naviguer dans une classification des pages indexées en catégories, puis formuler une requête sur le contenu de la catégorie correspondant le plus précisément à l'information recherchée. C'est également le type de stratégie proposé par les Online Public Access Catalogues (OPAC) permettant l'accès en ligne au contenu des bibliothèques américaines, ou par d'autres systèmes d'accès à d'importantes bases documentaires ([Dunlop & van Rijsbergen 93]). C'est également la démarche proposée dans la version du Livre de Connaissances Électronique de [Chaillot & Ermine 97] (voir chapitre 3), ainsi que dans d'autres systèmes d'accès à l'information à base de connaissances.

Inversement, en tant qu'utilisateur d'un moteur de recherche, un individu peut avoir la perception que l'une des réponses renvoyées est proche de l'information recherchée. Dans ce cas, il peut utiliser cette réponse comme « base de navigation » pour accéder à la solution de son problème d'information. La mise en œuvre de cette approche nécessite l'interrogation des nœuds d'un hypertexte (cf. [Balpe & al. 96], pp. 82-94). Par exemple, Y. Chiamarella & A.

Kheirbek construisent dynamiquement des « hyperindex » en fonction des requêtes des utilisateurs et proposent la navigation à partir de ces nœuds de l'index ([Chiamarella & Kheirbek 96]).

2.2.8. Les stratégies hybrides et les outils associés

D'autres stratégies d'accès à l'information ont été développées par hybridation des stratégies de browsing et de searching. Deux approches ont été envisagées. Il s'agit soit de calculs dynamiques de liens hypermédias à partir des résultats de requêtes, soit de formulation dynamique de requêtes à partir d'éléments d'hypertextes.

De nombreux travaux concernent la construction dynamique d'hypertextes grâce à des techniques empruntées à la recherche d'information. Plusieurs auteurs appliquent le modèle vectoriel de [Salton & McGill 83] pour la recherche d'information à la génération dynamique de liens hypermédias (cf. [Balpe & al. 96] pp. 101-105). Dans ce contexte, Jacques Savoy évalue la pertinence de plusieurs modèles en recherche d'information pour la construction d'hypertextes ([Savoy 96]). Nous nous focalisons ici sur des travaux permettant la construction de graphes de termes, qui nous semblent une approche représentative de ce type de systèmes.

Comme pour les systèmes présentés au paragraphe précédent, [Simoni et Fluhr 97], proposent un système à deux niveaux dans lequel l'index conceptuel est un graphe de termes construit automatiquement par des mécanismes d'extraction terminologique. Ces auteurs s'appuient sur le moteur d'indexation et de recherche SPIRIT qui permet cette hiérarchisation des termes à la fois pour la base documentaire et pour les requêtes. L'utilisateur navigue alors dans un le graphe des termes de sa requête ou directement dans la structure documentaire en fonction des liens calculés.

La démarche inverse a également été expérimentée. Les exemples ci-dessous concernent des systèmes hypermédias relatifs à la génération graphique de requêtes documentaires. La liste de ces exemples n'est pas exhaustive (cf. les « dynamic queries » de [Ahlberg & al. 92] ou le système WING de [Masui & al. 96] et Cat-a-Cone de [Hearst & Karadi 97], pour d'autres exemples de cette approche). Seuls trois systèmes y sont présentés, pour leurs spécificités par rapport à cette approche d'hybridation des stratégies de browsing et de searching.

Le système MORE, de Dario Lucarella et alii ([Lucarella & al. 93], [Lucarella & Zanzi 96a], [Lucarella & Zanzi 96b]), propose un modèle objet fondé sur la présentation et la manipulation graphiques d'éléments d'information multimédia. Ce type d'interaction est

également utilisé par les utilisateurs pour la formulation de requêtes, qui sont ensuite traitées par appariement des objets graphiques. Les définitions formelles de l'interrogation et du browsing dans cette collection d'éléments mettent l'accent sur les similarités de ces deux stratégies.

Les « Interactive Dynamic Maps » (IDM) de [Zizi & Beaudoin-Lafon 94] proposent deux niveaux de description contextuelle des documents. Ces IDM sont implémentées dans le système SHADOC qui permet la navigation dans une collection de documents. Des requêtes peuvent être générées graphiquement à partir des nœuds des IDM et sont formalisées en utilisant un modèle vectoriel des descripteurs des documents.

La stratégie d'« interactive querying », proposée par Gene Golovchinsky et Mark Chignell ([Golovchinsky & Chignell 93]), dans leur système « Queries-R-links » permettent la formulation de requêtes booléennes simples via la sélection graphique de mots dans un hyperdocument textuel. L'utilisateur indique alors graphiquement le type d'opérateur booléen qu'il souhaite utiliser pour relier les termes sélectionnés (seuls OU et ET sont disponibles). Ce système correspond à la stratégie hybride la plus avancée de cet état de l'art, puisqu'elle permet la construction de requêtes complexes à partir d'éléments graphiques d'un hypermédia.

2.2.9. *Bilan de l'accès à l'information*

Dans cette partie, nous avons dressé un état de l'art des stratégies et systèmes électroniques d'accès à l'information. Il ressort que les deux paradigmes de question-réponse pour la recherche d'information et de navigation pour les hypermédias peuvent être fusionnés selon plusieurs approches. Dans tous les cas, il s'agit de cumuler les avantages des deux stratégies, en prenant soin toutefois de s'affranchir de leurs inconvénients.

En particulier, il s'agit d'éviter les problèmes de désorientation posés par l'hypertextualisation des contenus, grâce à une mise en évidence de la structure conceptuelle des espaces de navigation. Les facteurs qui favorisent cette mise en évidence sont d'une part l'explicitation de la sémantique des éléments (nœuds, liens et ancres) de l'hypertexte et d'autre part la présentation d'un référentiel de navigation permettant à l'utilisateur de se situer dans cette structure. L'hybridation des deux types de stratégies d'accès à l'information permet en outre de s'affranchir de la complexité des langages de requêtes des outils d'interrogation.

La majorité des systèmes analysés ci-dessus comportent le plus souvent une couche documentaire et une couche d'index. Les systèmes d'accès à l'information décrits font

notamment appel à l'indexation conceptuelle des documents pour générer des graphes, le plus souvent arborescents, que l'utilisateur peut interroger ou parcourir.

2.3. Différentes approches d'aide à la compréhension de l'information

Une fois cette information accédée, il convient pour l'individu d'en prendre connaissance et de l'intégrer dans son « système de connaissances » (voir plus loin) en tant qu'élément de réponse au problème d'information initial (l'état anomal de connaissance). C'est ce que nous appellerons ici le processus d'interprétation ou de compréhension³⁴ de l'information.

Ce processus d'interprétation a été beaucoup étudié, et lui-même interprété dans le cadre de nombreuses approches telles que celles abordées ci-dessous. Cette énumération n'est pas exhaustive. Elle présente les grandes lignes des principales théories dont nous nous inspirons pour proposer des éléments de réponse à la problématique d'accès à l'information qui nous intéresse dans cette thèse.

Les deux facteurs déterminants de ce processus d'interprétation de l'information par les individus sont le facteur humain et l'information elle-même. Parmi les différentes approches présentées ci-dessous, les sciences cognitives concernent les aspects humains de ce processus, tandis que la sémiologie s'intéresse à la forme physique de l'information. De ces deux théories s'inspirent ce que nous pouvons qualifier, dans le cadre de notre problématique, de « disciplines d'aide à l'interprétation » de l'information. De telles disciplines sont la Gestion des Connaissances et l'Ergonomie des interfaces appliquée aux systèmes d'accès à l'information.

2.3.1. *L'approche cognitive*

En s'intéressant directement à l'état de connaissance d'un individu, les sciences cognitives se doivent de regrouper l'ensemble des disciplines qui ont pour but d'en expliquer le comportement. Ces différentes disciplines étudient donc, entre autres et sous des angles différents, cette problématique d'accès, de compréhension et d'interprétation de l'information. La neurologie aborde cette problématique sous l'angle des interactions électriques entre les différentes zones du cerveau lors du processus de communication, la linguistique sous celui de la communication orale entre les individus, et la sociologie s'y intéresse au niveau des groupes d'individus ([Ganascia 93], [Vignaux 94]). Ces disciplines sont relatives à des

³⁴ Au sens étymologique : prendre avec soi.

domaines d'expertise spécifiques, et il n'est pas dans nos objectifs ni dans nos compétences de les détailler toutes.

Nous ne nous intéresserons ici qu'à la branche la plus pertinente des sciences cognitives pour la problématique qui nous occupe, qui est, à notre avis, l'étude des processus d'apprentissage. Pour [Kintsch & Van Dijk 78] via [Dinet & al. 98], un processus d'apprentissage vise à la « production d'inférences nécessaires à l'établissement d'une représentation mentale stable et cohérente ». En d'autres termes, cette discipline s'emploie à modéliser le « système de connaissances d'un individu » à propos d'un domaine (celui-là même où existe le trou à l'origine de son problème d'information), et propose des méthodes pour faire évoluer ces modèles dans des directions données. Deux méthodes d'apprentissage sont classiquement distinguées en psychologie cognitive : l'apprentissage par l'instruction (ou didactique) et par l'action.

Dans le contexte d'application de cette discipline à la conception d'hypermédias, Jean-François Rouet et André Tricot ont identifié une troisième méthode d'apprentissage, baptisée « apprentissage par l'exploration » ([Rouet & Tricot 98]). Cette méthode présente l'avantage de se situer dans notre domaine d'application. De plus, d'un point de vue opérationnel, elle coïncide avec le modèle EST (évaluation, sélection, traitement) de ces auteurs.

2.3.2. L'approche sémiologique et sémiotique

Pour les sémiologues, le *signifiant* est l'information en tant que support physique de la communication³⁵, qui permet de véhiculer les connaissances entre les individus. Le *signifié* est le sens associé à ce signifiant chez l'individu qui l'émet ou le reçoit ([de Saussure 16], [Barthes 64]). Cependant, ce signifié émis ou produit est très dépendant de l'« état de connaissance » de l'individu chez qui il est considéré.

En sémiotique, le sens (interprétant) associé au signe (représentamen) l'est d'ailleurs par l'intermédiaire d'une relation triadique qui fait intervenir un signifiant (objet). Pierce définit trois relations qui relient les éléments de ce triangle sémiotique (Figure ci-dessous) : la syntactique, qui relie les signes entre eux, la sémantique, qui étudie la relation entre un signe et son signifié et la pragmatique, qui étudie l'utilisation qui est faite du signe ([Pierce 78] via [Peeters & Charlier 95]).

³⁵ Au sens de Shannon, ou encore l'*objet informationnel* (information-as-thing) de R. B. Allen (voir « définitions » au chapitre suivant). Ce concept englobe bien plus que le simple support papier ou numérique sur

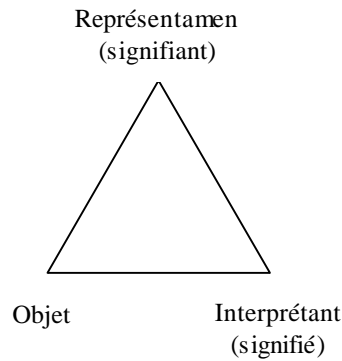


Figure 6. Le triangle sémiotique de Pierce, qui remplace la relation dyadique entre signifiant et signifié des sémiologues par une relation triadique faisant intervenir l'objet.

Cependant, malgré les tentatives de vulgarisation de la pensée peircienne, la nature de cet objet reste difficilement perceptible. C'est pourquoi nous nous référons à la psychologie cognitive, où la relation sémantique entre l'information et sa signification (le *bedeutung* de Frege) a été étudiée. Il ressort des différents travaux menés que cette signification du signe « repose initialement sur la prise en considération des événements psychologiques qui ont lieu dans l'intellect d'un individu, ou des structures cognitives qui s'y trouvent. Elle part donc de données individualisées. Toutefois, elle peut ensuite prendre en considération des entités qui sont communes à une multiplicité d'individus : par exemple, à tous les individus qui partagent la même langue » ([Peeters & Charlier 95]). Il apparaît donc que le sens associé à l'information est le résultat d'un processus complexe qui le rend inaccessible. Cependant, cette définition met également en évidence le fait que le contexte individuel ou social est un élément déterminant de ce processus.

Dans la mesure où le type d'information qui nous intéresse est l'information scientifique et technique, nous ne considérons ici comme déterminant pour le processus de communication que le contexte professionnel des individus³⁶. Ramené à l'information scientifique et technique, les trois éléments de ce triangle sont alors l'information, le contexte (i.e. l'activité professionnelle) dans lequel elle est émise ou reçue et le sens qui y est associé par l'individu émetteur ou récepteur de cette information. [Chaillot & Ermine 97] proposent le diagramme suivant pour illustrer le processus de communication entre les individus.

lequel est stockée l'information textuelle ou multimédia. Il désigne également la manifestation physique du signe auquel est associé un sens. Un feu rouge, une pendule ou un tableau sont alors de tels objets informationnels.

³⁶ Voir chapitre 3, section « Définitions », paragraphe « Connaissance ».

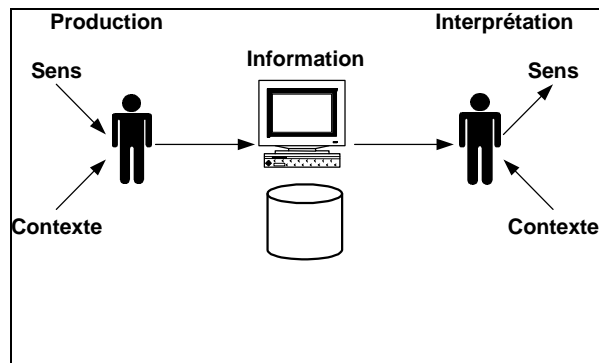


Figure 7. La transmission de l'information d'après [Chaillot & Ermine 97].

C'est pourquoi nous formulons l'hypothèse que la restitution du contexte d'émission d'une information peut aider l'individu la recevant à percevoir le sens qui lui a été associé lors de son émission (Figure 1). Il est alors utile pour l'utilisateur d'un système d'accès à l'information de disposer d'une représentation de ce contexte professionnel.

2.3.3. La gestion des connaissances

La Gestion des Connaissances est une discipline plus pragmatique qui s'inspire des sciences cognitives, humaines et de l'information. Historiquement, elle fait suite aux systèmes experts et à l'Intelligence Artificielle, mais se veut moins opérationnalisable que ces dernières disciplines. La gestion des connaissances définit des « flux de connaissances » circulant entre des individus ou des organisations ([Ermine 96], [Nonaka 95]). Elle s'emploie à les modéliser et à définir des méthodes de facilitation de transmission de ces connaissances. Pour cela, beaucoup de méthodes de gestion des connaissances s'appuient sur la systémique ([Von Bertalanffy 68], [de Rosnay 75], [Le Moigne 90]).

La gestion des connaissances définit des méthodes de modélisation des connaissances des experts d'un domaine. Ces méthodes visent à produire des livrables appelés « code de bonne pratique », « Manuel de Savoir-Faire » ou « Livre de Connaissances ». En fournissant un cadre méthodologique qui permet de formaliser les connaissances *déclaratives* et *procédurales* ([Ermine 96])³⁷ mises en œuvre dans une activité professionnelle, elles ont prouvé qu'elles peuvent effectivement permettre d'améliorer la communication interne ou la formation des nouveaux arrivants dans les entreprises.

En particulier, la méthode MKSM (Method for Knowledge Systems Management), développée à la DIST du CEA, restitue les différents points de vue liés à une activité autour

³⁷ Qui désignent en langage courant les « taxèmes » et « actèmes » de [Vogel 88].

d'un *système de connaissances* donné ([Ermine & al. 96]) en s'appuyant sur plusieurs formalismes graphiques (SADT, MAD, SCFC, réseaux sémantiques, arbres fonctionnels). Les *diagrammes de connaissances* sont élaborés à partir d'entretiens avec les experts du domaine et regroupés en un *Livre de Connaissances*. La richesse des diagrammes et des relations existant entre ces entités permet au lecteur d'appréhender la complexité du sujet traité. Cette méthode est décrite en détail au chapitre 3.

Toutefois, il nous paraît périlleux d'affirmer que de telles méthodes peuvent permettre de transmettre directement les connaissances d'un individu à un autre. En effet, que ce soit pendant la phase de modélisation ou d'utilisation du livrable produit, le « transfert » se fait sous forme d'information. Par conséquent, les connaissances modélisées suivent le même processus de production et d'interprétation que l'information et leur transmission est donc sujette aux mêmes aléas. C'est pourquoi nous préférons admettre, comme [de Saussure 16] que les connaissances au sens de la sémantique associée à l'information sont inaccessibles et différentes pour chaque individu.

Cependant, l'intérêt de la gestion des connaissances pour l'accès à l'information n'est pas négligeable. En effet, elle fournit un complément d'information très utile sur le contexte d'émission de l'information, en permettant de positionner les différentes notions les unes par rapport aux autres. La figure suivante illustre l'apport de cette transmission d'une représentation du contexte pour amélioration du processus de compréhension de l'information accédée.

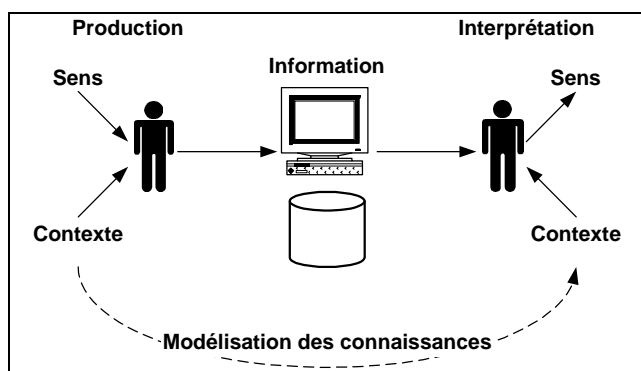


Figure 8. Intérêt de la Gestion des Connaissances dans le processus de transmission de l'information

Dans ce qui suit, nous considérons donc la Gestion des Connaissances comme cadre méthodologique nous permettant de modéliser et de formaliser le contexte d'émission de l'information échangée par les individus dans une organisation. Moyennant l'assumption que, dans un contexte similaire, le processus d'interprétation de l'information aboutit au même

« état de connaissance » que le processus d'émission de cette même information, nous pouvons considérer que la Gestion des Connaissances peut apporter une contribution efficace à la résolution des problèmes d'interprétation de l'information.

2.3.4. L'approche ergonomique comme moyen de contrôle des outils d'accès à l'information

Nous avons vu que tout processus de communication induit la circulation d'un flux physique d'information (les données), qui est véhiculé par un vecteur de communication (l'air, la lumière, une ligne téléphonique, l'Internet...). L'information est codée sous forme de données à la source de la communication, puis décodée une fois arrivée à destination. Ces opérations de codage et de décodage sont réalisées soit par l'individu lui-même, soit par des outils qui lui permettent d'étendre ses capacités de communication dans l'espace ou le temps. De tels outils sont par exemple le microphone et le haut-parleur, le clavier et l'écran (ou l'imprimante), les logiciels de traitement de texte et de visualisation ou encore un atelier de conception de livres de connaissances et un environnement de visualisation de ces livres de connaissances. Dans le cas d'une transmission non synchrone de l'information, des outils de stockage des données correspondantes entrent également dans le processus de communication.

Tous ces outils peuvent perturber ce processus de transmission de l'information pour les opérateurs humains qui le mettent en œuvre. Ils peuvent d'une part modifier l'information véhiculée³⁸. Les outils de communication peuvent également en provoquer la déformation, et donc la « mécompréhension »³⁹. Dans ces deux cas, le fonctionnement interne des outils d'aide à la communication (i.e. la transmission des données) n'est pas mis en cause. Cependant, ils provoquent des dysfonctionnements des processus de communication. Ces dysfonctionnements résultent d'une inadéquation de ces outils à l'opérateur humain, qui se manifestent lors de l'interaction d'un individu avec l'*interface*⁴⁰ d'un de ces systèmes.

³⁸ C'est par exemple le cas lorsqu'un individu qui n'est pas familiarisé avec l'usage d'un clavier adapte son texte pour que celui-ci soit plus court ou lorsque la taille maximale du message est dictée par l'outil. Un autre exemple courant est le phénomène de silence lié à l'utilisation d'un moteur de recherche d'information.

³⁹ Ce type de biais est par exemple courant en communication orale en environnement bruyé, lorsqu'un individu entend un mot à la place d'un autre. Nous assimilons ce phénomène à une mauvaise interprétation du signal sonore, de la même manière qu'une information textuelle lue sur un écran peut être mal interprétée lorsque cet écran est surchargé en informations non significatives.

⁴⁰ Ici, le terme d'interface d'un outil est à prendre au sens informatique, c'est-à-dire qu'il désigne le comportement des entrées et sorties de cet outil en provenance ou en direction de l'utilisateur.

L'ergonomie permet d'étudier et de faciliter l'interaction avec ces outils. En particulier, dans le cas d'outils informatiques, l'ergonomie des logiciels définit un cadre méthodologique pour la conception de ces outils et l'évaluation de l'*utilisabilité* de leur interface.

2.3.4.1. Les méthodes de modélisation et d'analyse de tâches

En amont de la conception, durant la phase de spécifications d'un système, l'ergonomie propose plusieurs types de méthodes de modélisation. Toutes ces méthodes sont fondées sur un recueil de données préalable concernant l'environnement et les méthodes de réalisation des tâches que la conception du système vise à suppléer. À partir de ces données, l'ergonomie propose des méthodes prédictives fondées sur la modélisation, comme GOMS (Goals, Operators, Methods, Selection, rules, cf. [John & Kieras 94]), HTA (Hierarchical Tasks Analysis, cf. [Shepherd 89]) ou MAD (Méthode Analytique de Description des tâches, cf. [Scapin 90]) qui permettent de modéliser les différentes tâches que les *utilisateurs finaux*⁴¹ vont avoir à réaliser avec le futur système et d'en déduire des fonctionnalités et une interface appropriées.

Dans le domaine de la conception de systèmes d'accès à l'information, la modélisation des tâches n'a été que partiellement étudiée. En particulier, la littérature est assez pauvre quant à l'analyse des tâches relatives aux stratégies de querying. Cela est sans doute dû à la distance qui existe entre la conception des outils d'accès à l'information utilisant ces stratégies (i.e. les moteurs de recherche) et celle des interfaces permettant d'y accéder. Dans ce domaine, la modélisation des tâches est alors plutôt le fait des concepteurs de l'interface que de ceux du moteur de recherche⁴².

En revanche, un plus grand nombre de travaux s'intéressent à l'analyse des tâches de navigation pour l'accès à l'information. [Tricot & Nanard 98] continuent cependant de souligner l'absence d'un cadre d'analyse spécifique à ce type d'applications. D'après ces auteurs, il semble toutefois que des formalismes tels que MAD permettent d'arriver à un bon niveau de description des tâches pour la spécification de l'interface. Il ne s'agit alors pas de décrire les tâches effectives de l'utilisateur du système, mais les tâches prescrites (ou procédure optimale) d'accès à l'information d'après les concepteurs du système. Ces auteurs

⁴¹ L'utilisateur final est la « cible » pour laquelle est conçu le système. Il est à opposer aux autres utilisateurs intermédiaires qui peuvent avoir à interagir avec le système pour le tester, avant sa validation et son installation définitive. La construction par dichotomie de cette locution est à rapprocher de celle la locution « indexation structurelle » évoquée plus haut.

⁴² Ce qui n'est pas nécessairement sans répercussion sur l'utilisabilité de l'interface, notamment du fait du langage de requêtes utilisé.

soulignent cependant que l'écart entre les tâches prescrites et effectives croît avec la complexité de la tâche à réaliser. Cependant, ils notent également « le cas des problèmes analogues qui entraînent des performances différentes ». Pour ces auteurs, cela est entre autres dû à l'« habillage verbal des problèmes ». D'après la compréhension que nous en avons, cette dernière notion est en rapport avec la mise en forme du problème. Dans le cadre de l'activité professionnelle des individus, on peut alors s'attendre à une certaine constance de cet habillage verbal, dû notamment à la formation commune des individus effectuant le même travail. De la même manière, nous pouvons supposer que la formation à un outil d'accès à l'information permet aux différents individus de se faire une représentation du problème commune, et si possible relativement proche de celle de la tâche prescrite par le concepteur.

Plus proches du domaine de la psychologie cognitive, les travaux de [Rouet & Tricot 98] définissent un modèle des processus impliqués dans la tâche de recherche d'information dans un hypertexte. Ce modèle en trois étapes (évaluation, sélection, traitement) permet de décrire l'activité de navigation de l'utilisateur. Pour ses auteurs ce modèle contraste avec l'approche formelle de la modélisation des tâches en ce qu'il prend en compte « la capacité de traitement limité de la MDT » (i.e. la mémoire de travail) de l'opérateur humain⁴³. Pour ces auteurs, ce modèle doit permettre aux concepteurs de systèmes hypermédias d'accès à l'information d'améliorer les qualités ergonomiques de ces systèmes en facilitant, pour chaque étape de la navigation, les tâches d'évaluation, de sélection et de traitement de l'information. Cela peut être fait en offrant à l'utilisateur la possibilité de visualiser l'information disponible, en lui permettant un accès intelligible aux différentes étapes de la navigation et enfin en optimisant la lisibilité du contenu.

2.3.4.2. Les profils d'utilisateurs

La modélisation des utilisateurs est, au même titre que celle des tâches, une donnée essentielle pour la conception ([Allen 97], [Palmquist 98]). Elle peut passer par la mise au point de *profils d'utilisateurs* spécifiques, dans le cas où le concepteur sait qu'il aura affaire à une population d'utilisateurs hétérogène. En effet, pour [Waern & Hägglund 97], « to be able to meet the psychological requirements, an explanation must be relevant to the knowledge of the recipient. Thus, a User Model is necessary. » Cet article fait, entre autres, référence à [Kass & Finin 88] pour justifier cette affirmation. En particulier, une synthèse des travaux réalisés sur

⁴³ Cette capacité limitée a pour principale conséquence le phénomène de désorientation décrit plus haut.

les différences de comportements entre utilisateurs novices et experts d'un système est présentée dans [Mayer 97].

Dans le cas des systèmes d'accès à l'information, des exemples de systèmes ont été conçus en intégrant des profils des utilisateurs. Le but le plus communément rencontré de ces profils est la limitation du bruit (nombre de documents non pertinents retournés par le système) grâce à un filtrage des réponses en fonction des entrées d'intérêt connus de l'utilisateur. Une autre application est la démarche opposée : il s'agit d'élargir la requête (« query expansion ») de façon à limiter le silence (nombre de documents pertinents non retournés par le système). Dans tous les cas, ces profils identifient plusieurs caractéristiques de l'utilisateur. Une étude détaillée de ces caractéristiques spécifiques à la problématique de cette thèse est présentée au chapitre 3.

2.3.4.3. Les règles de conception

Pendant la phase de développement, l'ergonomie intervient en proposant une série de règles de conception, qui permettent, sinon d'assurer la parfaite utilisabilité de l'interface, du moins d'en éviter les principales erreurs de conception. Il existe des règles de conception générales ([Normann 86], [Schneiderman 92], [Mayhew 92], [Smith 96]) et spécifiquement adaptées au type de système à concevoir.

Dans le cas des outils d'accès à l'information dont nous avons parlé plus haut, l'ergonomie est d'autant plus importante que ces systèmes ont pour but de restituer une information préexistante, sans y apporter aucun traitement. Leur raison d'être est donc de permettre et de faciliter cette restitution à l'utilisateur. Cependant, en fonction du type de système d'accès à l'information, les méthodes de conception sont différentes : un moteur d'indexation et de recherche n'a a priori pas de contact avec l'utilisateur ; il lui faut justement un autre programme qui se charge de l'interface ; pour les systèmes hypermédia, qui sont entièrement déterminés par leurs interfaces, il existe des règles de conception ergonomiques spécifiques ([Nielsen 90] pp. 163-171, [Woodhead 90], pp. 97-150, [Mc Knight & al. 91], pp. 87-104, [Balpe & al. 96], [Nanard & Nanard 98]). Parmi celles-ci⁴⁴, des règles et méthodes de conception d'hypermédiats ont été formulées spécifiquement pour les hypermédiats dédiés à la recherche d'information ([Lucarella 90], [Rouet & Tricot 95], [Rouet & Tricot 98], voir paragraphe précédent) ou à l'apprentissage ([Hammond & Allinson 89], [Tricot & al. 98]). De nombreux travaux concernent également les hypermédiats dédiés à l'Internet ou à un Intranet

(nous ne citons ici que l'étude, complète et en français, de [Bastien & al. 98], mais un historique et une synthèse de ces travaux est disponible – en anglais – dans [Vora & Helander 97]).

2.3.4.4. Les méthodes d'évaluation

Une fois le système réalisé, l'ergonomie met à la disposition des concepteurs plusieurs méthodes de validation. Une synthèse des méthodes existantes est présentée dans [Farenc 00]. Cet auteur identifie plusieurs distinctions entre les catégories de méthodes⁴⁵ d'évaluation : prédictives vs. expérimentales, empiriques vs. analytiques ou fondées sur l'utilisateur vs. fondées sur la tâche. Les méthodes prédictives concernent les règles de conception décrites au paragraphe précédent. Nous ne nous intéressons ici qu'aux méthodes expérimentales. Nous revenons plus loin sur les caractéristiques des autres distinctions.

Dans le cas des systèmes hypermédias d'accès à l'information, deux types de buts sont à prendre en compte : il s'agit de mesurer d'une part l'utilité du système (l'efficacité de la recherche d'information, en termes de bruit/silence ou de rappel/précision) et d'autre part l'utilisabilité de l'interface de l'hypermédia. Les méthodes d'évaluation de la pertinence des documents retrouvés par le système correspondent aux évaluations analytiques de Christelle Farenc et ont été décrites précédemment.

Concernant l'évaluation de systèmes hypermédias, un « Panel » animé par Gary Perlman propose un état de l'art de ces méthodes [Perlmann & al. 90]. Il ressort de cet état de l'art que les méthodes d'évaluations doivent porter sur les caractéristiques problématiques de ces systèmes identifiées précédemment. Il s'agit d'évaluer (en mesurant par exemple le nombre de clicks nécessaires pour l'accès à l'information) pertinence d'une stratégie d'accès à l'information. L'analyse du cheminement parcouru permet de déterminer si un utilisateur est désorienté. La « compréhensibilité » de la structuration de l'hypertexte n'est pas abordée dans cet article. Nous attribuons cela au fait que cette session s'est tenue avant les travaux évoqués plus haut sur cet aspect des systèmes hypertextes.

Plus récemment, et dans le cadre des évaluations de sites web, [Bastien & al. 98] définissent quatre axes pour l'évaluation ergonomiques des hypermédias : les tests d'évaluation, les

⁴⁴ Nous ne donnons ici que des exemples concernant les types de systèmes qui nous sont utiles dans cette thèse.

⁴⁵ Nous employons ici ce terme dans son sens générique, contrairement à la distinction faite par C. Farenc, qui les considère comme des techniques, par opposition aux outils conceptuels que sont les méthodes « académiques » de modélisation, telles que GOMS ou MAD.

questionnaires et entretiens, les méthodes d'inspection ergonomiques et l'évaluation automatique.

- Les tests d'évaluation correspondent aux méthodes empiriques de [Farenc 00]. Il s'agit de tests menés auprès d'utilisateurs du système, que ceux-ci soit les réels destinataires de ce système ou non⁴⁶. Ces tests peuvent être menés en environnement contrôlé ou sur site. Un environnement contrôlé est un laboratoire spécifique, dans lequel l'utilisateur est isolé pour interagir avec le système. Il est alors observé et ses réactions sont notées et analysées. Ce type d'environnement permet un gain de temps et de coût dans l'évaluation des interfaces, car l'isolement permet d'éliminer toute interférence de l'extérieur sur cette interaction, et donc d'observer directement le comportement de l'utilisateur face au système. Le site est l'environnement de travail réel des utilisateurs finaux. L'observation des utilisateurs sur site pose souvent des problèmes logistiques et humains, mais elle permet d'évaluer l'utilisabilité du système dans des conditions réelles d'utilisation (cf. « situation theory », [Cooper & al.] via [Huibers & al 96]). L'intérêt des tests d'évaluation est de permettre d'obtenir des données quantitatives et qualitatives sur les performances réelles d'un système (temps, nombre de clicks nécessaires à l'utilisateur pour réaliser une tâche, problèmes et erreurs rencontrés...). Le principal désavantage de ces méthodes est leur coût, puisqu'elles nécessitent la présence à la fois des évaluateurs et des utilisateurs finaux. Comme nous l'avons signalé plus haut, le coût du matériel spécifique à l'évaluation et de la logistique entre également en ligne de compte.

- Les questionnaires et les entretiens permettent aux utilisateurs de donner leur avis a posteriori sur l'utilité du système. Ils sont moins coûteux à mettre en place, mais donnent parfois des résultats subjectifs. Cette subjectivité est cependant compensée en obtenant des données provenant d'un grand nombre d'utilisateurs. S'il aborde les différentes tâches à réaliser par les utilisateurs à un niveau de finesse suffisant, un questionnaire permet d'obtenir une évaluation détaillée des fonctionnalités correspondantes. Un exemple de ce type de questionnaires (le questionnaire WAMMI) est donné en page 153 de [Bastien & al. 98]).

- Les méthodes d'inspection ergonomique permettent d'évaluer l'utilisabilité du système et son adéquation aux règles de conception ergonomiques ou à des « normes, principes, dimensions, heuristiques » ([Bastien & al. 98]). En pratique, il s'agit, pour les évaluateurs, de valider la conformité du système par rapport à des critères ergonomiques d'évaluation, tels que ceux proposés dans [Bastien & Scapin 93].

- L'évaluation automatique s'appuie sur des outils capables de simuler le parcours des sites par un utilisateur. Les performances de ces outils sont cependant limitées, et la totalité des critères ergonomiques à évaluer n'est pas prise en compte.

Pour être complet, nous devons signaler l'existence de solutions hybridant inspection et tests utilisateurs. C'est le cas de la méthode SUE (Systematic Usability Evaluation), proposée par [Garzotto & Matera 97, (via [Garzotto & al. 99]). Cette méthode, qui n'est pas spécifique aux hypermédias, propose un cadre méthodologique général pour l'évaluation de systèmes interactifs. Elle fait appel à une inspection ergonomique pour définir des « concrete tasks », qui sont ensuite évaluées grâce à des tests utilisateurs. L'intérêt de cette méthode est de permettre l'économie de tests utilisateurs pour des tâches qui n'ont pas révélé de problèmes particuliers lors de l'inspection ergonomique. Elle se limite toutefois à une évaluation de l'utilisabilité du système, sans s'intéresser à son utilité. Appliquée au modèle hypermédia générique HDM (Hypermedia Design Model, cf. [Garzotto & al. 93]), cette méthode permet la définition d'« Abstract Tasks » (AT) spécifiques aux hypermédias ([Garzotto & al. 99]). Ces AT peuvent alors fournir une méthodologie d'évaluation ergonomique systématique de l'utilisabilité d'applications hypermédia. L'avantage de ces méthodes (i.e. SUE et les AT) est une certaine indépendance des résultats vis-à-vis de l'expérience de l'évaluateur et une évaluation plus détaillée que les méthodes automatiques. En contrepartie, elles requièrent l'intervention d'un opérateur humain, et nécessitent un modèle spécifique (HDM pour les AT) de l'application à évaluer.

Il apparaît que ces méthodes hybrides permettent un gain de temps et limitent le coût des campagnes d'évaluation de l'utilisabilité au minimum. Cependant, l'évaluation de l'utilité du système doit s'appuyer sur des retours de la part des utilisateurs. Il nous semble donc qu'une démarche d'évaluation rationnelle peut débuter par une méthode d'évaluation de l'utilisabilité ne nécessitant pas la présence d'utilisateurs, puis se poursuivre par une validation de cette utilisabilité et une évaluation de l'utilité du système auprès des utilisateurs. Pour limiter les coûts, un recours limité à des tests utilisateurs pourrait être allié à un questionnaire.

Par conséquent, nous voyons que l'ergonomie des logiciels influe sur le processus de transmission et d'accès à l'information assistée par ordinateur. L'intérêt d'une démarche de conception ergonomique de ce type de systèmes se comprend dans le but concevoir un système utilisable et utile permettant de mettre facilement en œuvre une stratégie d'accès à

⁴⁶ Dans le cas où les tests sont menés auprès des utilisateurs finaux et où plusieurs catégories d'utilisateurs sont identifiées, il convient de choisir les utilisateurs testeurs de façon représentative dans ces différentes catégories.

l'information adaptée au niveau du problème d'information. Sa validation doit nécessairement comporter au moins deux phases d'évaluation.

2.4. Bilan de l'état de l'art de la problématique de l'accès à l'information

Dans ce chapitre, nous avons identifié trois composantes de l'accès à l'information. Il apparaît tout d'abord qu'une démarche d'accès à l'information est la conséquence de la perception par un individu d'un problème d'information. Celui-ci peut avoir plusieurs niveaux de précision. L'accès à l'information relative à un problème d'information au niveau ressenti relève des processus d'apprentissage. Nous considérons que l'accès à l'information dans le cadre de l'exercice d'une activité professionnelle a lieu à partir d'un problème d'information perçu au niveau conscient. Cette notion de problème d'information, même si elle ne reflète pas les travaux les plus récents de la littérature concernant la démarche d'accès à l'information, fournit tout de même un cadre structurant que nous considérons comme suffisant pour l'approche de la problématique d'accès à l'information utilisée dans cette thèse.

Dans une deuxième partie, nous passons en revue les différentes stratégies d'accès à l'information et les outils associés. Il ressort de cet état de l'art que les deux principales stratégies (searching et browsing) et les outils associés (moteurs de recherche et hypermédia) peuvent être utilisés conjointement. En particulier, des outils spécifiques sont désormais développés pour permettre à l'utilisateur d'accéder à l'information grâce à des stratégies hybrides. L'avantage de ces stratégies est de permettre l'interrogation à un niveau conceptuel des corpus documentaires sans formulation explicite de requête. Cependant, les requêtes générées sont souvent limitées. Par exemple, l'interrogation de graphes de concepts ne permet que la spécification de relations d'inclusion, et les mécanismes de construction graphiques de requêtes permettent au mieux l'utilisation de deux opérateurs booléens (ET et OU).

La troisième partie de ce chapitre passe en revue plusieurs disciplines permettant, une fois l'information pertinente pour un problème identifiée et visualisée, d'en améliorer l'interprétation. Parmi celles-ci, la gestion des connaissances propose une approche par la recontextualisation qui nous semble mériter d'être explorée. L'ergonomie des logiciels nous apparaît également comme une composante essentielle de la phase de compréhension, en proposant notamment l'emploi de profils pour permettre la prise en compte des spécificités des utilisateurs. Cette discipline permet en outre de favoriser la cohérence entre la stratégie adoptée par les concepteurs des systèmes et les résultats de la mise en œuvre de cette stratégie

par les utilisateurs. Dans ce domaine, les travaux sur l'analyse des tâches révèlent un manque dans le domaine des systèmes électroniques d'accès à l'information. Cette lacune est attribuée à l'impossibilité de considérer l'accès à l'information comme une activité isolée. La recherche d'information ne se conçoit alors que par rapport au contexte spécifique de l'activité ayant provoqué l'apparition du problème d'information. Nous pensons que l'identification de ce contexte peut être utilement fournie par les méthodes de modélisation utilisées en gestion des connaissances.

Les trois disciplines balayées par cet état de l'art sont donc les sciences de l'information, la gestion des connaissances et l'ergonomie. Il nous semble qu'une approche pluridisciplinaire fondée sur la modélisation du contexte de l'accès à l'information et les stratégies hybrides d'accès à l'information peut favoriser les processus d'accès à l'information des individus dans l'accomplissement de leur activité professionnelle.

Bibliographie

- [ACM 92] : Association for Computer Machinery, Special Interest Group on Computer-Human Interaction, *"Curricula for Human-Computer Interaction"*, ISBN 0-89791-474-0, disponible en ligne : <http://www.acm.org/sigchi/cdg/cdg2.html>, 1992.
- [Agosti & al. 91] : Agosti M., Colotti R., Gradenigo G., *"A two-level hypertext retrieval model for legal data"*, in proceedings of the ACM Hypertext'91 Conference, pp. 316-325, 1991.
- [Agosti & Smeaton 96] : *"Information Retrieval and Hypertext"*, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, 1996.
- [Ahlberg & al. 92] : Ahlberg C., Williamson C., Shneiderman B., *"Dynamic Queries for Information Exploration: An Implementation and Evaluation"*, in proceedings of the ACM CHI'92 Conference on Computer-Human Interaction, pp. 619-626, Mai 1992.
- [Ahlswede 88] : Ahlswede T., *"Automatic construction of phrasal thesaurus for an information retrieval system form a machine readable dictionnary"*, RIAO 88, vol. 1, pp. 597-608, 1988.
- [Alix & Roussel 01] : Alix C., Roussel C., *"La vie est mouvante, c'est cette complexité qu'il faut représenter"*, interview de T. Nelson, in Week-end Rencontre, Libération, pp. 36-37, édition du 26 & 27 mai 2001.
- [Allen 97] : Allen R. B., *"Mental Models and User Models"*, Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu (eds.), Elsevier Science, 1997, pp.49-63.
- [Assadi & Gros 99] : Assadi H., Gros C., *"Ingénierie des connaissances pour les systèmes de consultation de documentation technique"*, Document Numérique, Hermès, vol. 3, n° 3-4, pp. 15-35, 1999.
- [Balpe & al. 96] : Balpe J-P., Lelu A., Papy F., Saleh I., *"Techniques avancées pour l'hypertexte"*, Hermès, Paris 1996.
- [Bastien & Scapin 93] : Bastien, J. M. C., Scapin D-L, *"Ergonomic Criteria for the Evaluation of Human-Computer Interfaces"*, Rapport technique INRIA n° 156, Juin 1993.
- [Bastien & al. 98] : Bastien, J. M. C., Leulier C., Scapin D-L, *"L'Ergonomie des sites web"*, in Créer et Maintenir un service web, J-C. Le Moal & B. Hidoine Eds., ADBS, Paris, pp. 111-173, 1998.
- [Barker 97] : Barker A. L, *"DataStar Web: living up to the hype? An evaluation of the interface and search system"*, in proceedings of the Online information'97 Conference, pp. 213-222, 1997.
- [Barthes 64] : Barthes R., *"Éléments de sémiologie"*, 1964.
- [Bates 89] : Bates M., *"The design of browsing and berrypicking techniques for the on-line search interface"*, Online Review, 13(5), pp. 407-431, 1989.
- [Bateson 87] : Bateson A. G., Alexander R. A., Murphy M. D., *"Cognitive processing differences between novice and expert computer programmers"*, International Journal of Man-Machine Studies, vol. 26, pp. 649-660, 1987.
- [Belkin 80] : Belkin N. J., *"Anomalous states of knowledge as a basis for Information retrieval"*, Canadian Journal of Information Science, n°5, pp. 133-143, 1980.

- [Belkin & al. 93] : Belkin N. J., Marchetti P. G., Cool Chen, *"BRAQUE: Design of an interface to support user interaction in information retrieval"*, Information Processing and Management, vol. 29, n° 3, pp. 325-344, 1993.
- [Bernstein 90] : Bernstein M., *"Hypertext and technical writing"*, in proceedings of the ECHT'90 Course, Versailles, Novembre 1990.
- [Blanger 98] : Blanger J-P., Notes personnelles prises lors de la formation au logiciel Acrobat, BVA MYFRA, octobre 1998.
- [Brusilovsky 96] : Brusilovsky P., *"Methods and techniques of adaptive hypermedia"*, User Modeling and User-Adapted Interaction, Special issue on adaptive hypertext and hypermedia, 6 (2-3), pp. 87-129, 1996, disponible en ligne : <http://citeseer.nj.nec.com/cache/papers2/cs/440/http:zSzzSzwww.contrib.andrew.cmu.edu/Sz~plbzSzUMUAI.pdf/brusilovsky96methods.pdf>
- [Bush 45] : Bush V., *"As we may think"*, Atlantic Monthly, 176 (1), pp. 101-108, 1945.
- [Chaillot & Ermine 97] : Chaillot M., Ermine J-L., *"Le livre de connaissances électronique"*, Document numérique, vol. 1, n° 1, 75-98, 1997.
- [Chiamarella & Kheirbek 96] : Chiamarella Y., Kheirbek A., *"An intergrated model for hypermedia and information retrieval"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 139-178, 1996.
- [Cooper & al. 1990] : Cooper R., Mukai K., Perry J., *"Situation Theory and its Applications"*, vol. 1, CSLI, 1990.
- [Crampes 99] : Crampes M., *Introduction à l'atelier sur les Documents Virtuels Personnalisables*, in Interaction Homme-Machine, Actes de l'Atelier Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, Novembre 1999.
- [de Saussure 16] : de Saussure F., *"Cours de linguistique générale"*, (s.l.), 1916.
- [de Rosnay 75] : de Rosnay, J., *"Le macroscope : vers une vision globale"*, Seuil, Paris, 1975.
- [Dervin & Nilan 86] : Dervin B., Nilan M., *"Information needs and uses"*, Annual Review of Information Science and Technology, M. Williams (Ed.), White Plains, New York: Knowledge Industries, vol. 21, No. 3, pp. 33, 1986.
- [Dinet & al. 98] : Dinet J., Rouet J-F, Passerault J-M., *"Les « Nouveaux outils » de recherche documentaire sont-ils compatibles avec les stratégies cognitives des élèves ?"*, in actes du 4^{ème} Colloque Hypermédias et Apprentissages, J-F Rouet et B. de la Passardière Eds., Poitiers, pp. 149-161, octobre 1998.
- [DISTNB 98] : Direction de l'information scientifique, des technologies nouvelles et des bibliothèques (DISTNB), Département de l'information spécialisée, Définition de l'IST disponible en ligne : <http://www.mesr.fr/icst/ist/index.htm>, 1998.
- [Dunlop & van Rijsbergen 93] : Dunlop M., Van Rijsbergen C. J., *"Hypermedia and free text retrieval"*, Information Processing and Management, vol 29, n° 3, pp. 287-298, 1993.
- [Eco 88] : Eco, U., *"Le signe"*, Livre de poche, Essais, 1988.
- [Engelbart 62] : Engelbart D. C., *"Augmenting human intellect: A conceptual framework"*, Technical Report AF 49(638)-1024, Stanford Research Institute, Octobre 1962.
- [Erdmann & Studer 00] : Erdmann M., Studer R., *"How to Structure and Access XML Documents With Ontologies"*, Data and Knowledge Engineering, special issue on

Intelligent Information Integration, 2000.

- [Ermine 89] : Ermine J-L, "*Systèmes experts, théorie et pratique*", Collection Tec et Doc, Lavoisier Ed, Paris, 1989.
- [Ermine 96] : Ermine J-L., "*Les systèmes de connaissances*", Hermès, 1996.
- [Ermine & al. 96] : Ermine J-L., M. Chaillot, P. Bignon, B. Charreton, D. Malavieille: "*MKSM, a method for knowledge systems management*". in Knowledge Management, Organization, Competence and Methodology, Advances in Knowledge Management, Volume 1, Jos F. Schreinmakers Ed., pp. 288-302, Ergon Verlag, 1996.
- [Farenc 00] : Farenc C., "*ERGOVAL : une méthode de structuration des règles ergonomiques permettant l'évaluation automatique d'interfaces graphiques*", Thèse de doctorat, Chapitre 1 – Les méthodes d'évaluation, version du 11 décembre 2000, disponible en ligne : <http://lis.univ-tlse1.fr/farenc/nouvelle.htm>.
- [Foucault 99] : Foucault B., "*Analyse des stratégies de navigation sur le Web : expériences sur les modes d'accès à l'information*", in Interaction Homme-Machine, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Cépaduès-Éditions, Toulouse, France, vol. 2, pp. 45-48, 1999.
- [Ganascia 93] : Ganascia J-G., "*Les Sciences Cognitives*", Collection Que Sais-Je ?, Dunod, Paris, 1993.
- [Garzotto & al. 91] : Garzotto F., Mainetti L., Paolini P., "*HDM: a model for the design of hypertext applications*", in proceedings of the ACM Hypertext'91 Conference, pp. 313-328, décembre 1991.
- [Garzotto & al. 93] : Garzotto F., Mainetti L., Paolini P., "*HDM: a model based approach to hypertext application design*", in proceedings of the ACM Transactions on Information Systems, vol. 11, n° 1, pp. 1-26, 1993.
- [Garzotto & al. 99] : Garzotto F., Matera M., Paolini P., "*Abstract Tasks: a Tool for the inspection of Web Sites and Off-line Hypermedia*", in proceedings of the tenth ACM Conference on Hypertext and hypermedia, pp. 157-163, 1999.
- [Golovchinsky & Chignell 93] : Golovchinsky G., Chignell M., "*Queries-R-Links: Graphical Markup for Text Navigation*", in proceedings of the ACM INTERCHI'93, Conference on Human Factors in Computing Systems, pp. 454-460, 1993.
- [Golovchinsky 97] : Golovchinsky, G., "*Queries? Links? Is there a difference?*", in proceedings of the ACM CHI'97 Conference on Human Factors in Computing Systems, vol. 1, pp. 407-414, 1997.
- [Guarino & al. 99] : Guarino N., Masolo C., Vetere G., "*OntoSeek: Content-Based Access to the Web*", in IEEE Intelligent Systems, pp. 70-80, Mai-Juin 1999.
- [Halasz & Schwartz 94] : Halasz F. G., Schwartz M., "*The Dexter hypertext reference model*", Communication of the ACM, vol. 37, n° 2, pp.30-39, 1994.
- [Hammond & Allinson 89] : Hammond N., Allinson L., "*Extending Hypermedia for Learning: An Investigation of Access and Guidance Tools*", in Proceedings of the HCI'89 Conference on People and Computers, pp. 293-304, 1989.
- [Hardman & al. 94] : Hardman L., Bulterman D., Van Rossum G., "*The Amsterdam hypermedia model*", Communication of the ACM, vol. 37, n° 2, pp. 50-63, 1994.
- [Hearst & Karadi 97] : Hearst M. A., Karadi C., "*Cat-a-Cone: An Interactive Interface for*

- Specifying Searches and Viewing Retrieval using Large Category Hierarchy*", in proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, p.246-255, 1997.
- [Heflin & al. 99] : Heflin, J., Hendler J., Luke S., *"SHOE: a Knowledge Representation Language for Internet Applications"*, Technical Report CS-TR-4078 (UMIACS TR-99-71), Dept of Computer Science, University of Maryland at College Park, 1999, disponible en ligne : <http://www.cs.umd.edu/projects/plus/SHOE/pubs/techrpt99.pdf>
- [Huibers & al. 96] : Huibers T., Ounis I., Chevallet J-P., *"Conceptual Graph Aboutness"*, in proceedings of the 4th International Conference on Conceptual Structures, ICCS'96, Lecture Notes in Artificial Intelligence, P. W. Eklund, G. ELLIS, G. Mann Eds., pp. 130-145, Sydney, 1996.
- [IGL 82] : IGL, France : *"Introduction à SADT"*, Manuel, 1982.
- [Ingwersen 94] : Ingwersen P., *"Polyrepresentation of information needs and semantics entities: elements of a cognitive theory for information retrieval interaction"*, in Proceedings of the seventh Annual International ACM-SIGIR, Dublin, Ireland, 1994.
- [Iwayama & Tokunaga 95] : Iwayama M., Tokunaga T., *"Cluster-Based Text Categorization: A Comparison of Category Search Strategies Text Categorization"*, in proceedings of the 8th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval p.273-280, 1995.
- [John & Kieras 94] : John, B. E. , Kieras, D. E., *"The GOMS family of analysis techniques : tools for design and evaluation"*, Carnegie Mellon University, Pittsburg, 1994.
- [Jolion 01] : *"L'indexation"*, numéro spécial de la revue Document Numérique, sous la direction de J-M Jolion, Document Numérique, Vol 4, n°1-2/2000, Hermès Science Publication, 2001.
- [Jouis, 93] : Jouis C., *"Contributions à la conceptualisation et à la modélisation des connaissances à partir d'une analyse linguistique de textes"*, Thèse de doctorat de l'École des Hautes Études en Sciences Sociales, 1993.
- [Kass & Finin 88] : Kass T., Finin T., *"The Need for User Models in Generating Expert System Explanations"*, in International Journal of Expert Systems Research and Applications, vol. 1., pp. 345-375, 1988.
- [Kintsch & Van Dijk 78] : Kintsch W., Van Dijk T. A., *"Towards a Model of Text Comprehension and Production"*, Psychological Review, vol. 85, n° 5, pp. 363-394, 1978.
- [Kolmayer 98] : Kolmayer E., *"Démarche d'interrogation documentaire et navigation"*, in Quatrième Colloque Hypermedias et Apprentissages, Poitiers, pp. 89-96, 15-17 octobre 1998.
- [Lainé-Cruzel 94] : Lainé-Cruzel S., *"Vers de nouveaux systèmes d'information prenant en compte le profil des utilisateurs"*, Documentaliste – Sciences de l'information, vol. 31, n° 3, pp. 143-147, 1994.
- [Lainé-Cruzel & al. 96] : Lainé-Cruzel S., Lafouge T., Lardy J-P. & Ben Abdallah N., *"Improving information retrieval by combining user profile and document segmentation"*, Information Processing and Management, Elsevier Science Ltd, Vol. 32, No. 3, pp. 305-315, 1996.
- [Le Loarer 1994] : Le Loarer P., *"Indexation automatique, recherche d'information et évaluation"*, Cours INRIA, Le traitement électronique du document, Aix-en-Provence, pp.

149-201, octobre 1994.

- [Leloup 98] Leloup, C., *"Moteurs d'indexation et de recherche, environnements client-serveur, internet et intranet"*, Eyrolles, Paris, 1998
- [Lelu 93] : Lelu A., *"Modèles neuronaux pour l'analyse de données documentaires et textuelles – organiser de très grands tableaux de données qualitatives en pôles et zones d'influence"*, Thèse de l'université Paris 6, 1993.
- [Lelu & al. 99] : Lelu A., Hallab M., Rhissassi H., Papy F., Bouyahi S., Bouhaï N., He H., Qi C., Saleh I., *"Projet NeuroWeb : un moteur de recherche multilingue et cartographique"*, actes de la 5^{ème} conférence internationale Hypertextes, Hypermédias et Internet, H²PTM'99, réalisations, outils et méthodes, Hermès, Paris, pp. 43-56, 1999.
- [Le Moigne 77] : Le Moigne, J-L., *"La théorie du système général, Théorie de la modélisation"*, Paris, PUF, 1977, 3^{ème} édition mise à jour, 1990.
- [Le Moigne 90] : Le Moigne, J-L., *"La modélisation des systèmes complexes"*, Dunod, 1990.
- [Lesk 89] : Lesk M., *"What to do when there's too much information"*, in proceedings of the ACM Hypertext'89 conference, pp. 305-318.
- [Lucarella 90] : Lucarella D., *"A Model for Hypertext-Based Information Retrieval. Formal Models and Query Languages"*, in Proceedings of the ECHT'90 European Conference on HyperTexts, pp. 81-94, 1990.
- [Lucarella & al. 93] : Lucarella D., Parisotto S., Zanzi A., *"MORE: Multimedia Object Retrieval Environment"*, in proceedings of the ACM Hypertext'93, pp. 39-50, 1996.
- [Lucarella & Zanzi 96a] : Lucarella D., Zanzi A., *"A Visual Retrieval Environment for Hypermedia Information Systems"*, in proceedings of the ACM Transaction on Information Systems, vol. 14, n° 1, pp. 3-29, 1996.
- [Lucarella & Zanzi 96b] : Lucarella D., Zanzi A., *"Information modelling and retrieval in hypermedia systems"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 121-138, 1996.
- [Mackay 60] : Mackay D. M., *"What makes the question"*, The listener, n°63, pp. 789-790, 1960.
- [Mackinlay & al. 95] Mackinlay J. D., Zellweger P. T., Chignell M., Furnas G., Salton G., *"Browsing vs. Search: Can We Find a Synergy?"*, Panels, in proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems, ACM Press, vol 2, pp. 179-180, 1995.
- [Masui & al. 96] : Masui T., Minakuchi M., Borden G. R., Kashiwagi K., *"WING: A Multiple-View Smooth Information Retrieval System"*, in proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, System Demonstrations, p. 338, 1996.
- [Marchionini & Shneiderman 88] : Marchionini G., Shneiderman B., *"Finding facts vs. browsing knowledge in hypertexts systems"*, IEEE computer, vol. 21, N° 1, pp. 70-80, 1988.
- [Marchionini 89] : Marchionini G., *"Information seeking in electronic encyclopaedia"*, Machine-Mediated Learning, vol. 3, N° 3, pp. 211-226, 1989.
- [Marchionini 95] : Marchionini G., *"Information seeking in electronic environments"*, New York, Cambridge University Press, 1995.

- [Martin & Eklund 00] : Martin P., Eklund P., *"Knowledge Retrieval and the Word Wide Web"*, in IEEE Intelligent Systems, Special Issue on Knowledge Management and the Internet, Mai-Juin 2000.
- [Masand et al., 1992] : Masand B., Linoff G., Waltz D., *"Classifying news stories using memory based reasoning"*, in proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 59-65, 1992.
- [Mayer 97] : Mayer. R. E., *"From Novice to Expert"*, Handbook of Human-Computer Interaction, M. Helander, T. K. Landhauer, P. Prabhu eds., Elsevier Science B.V., pp.781-795, 1997.
- [Mayhew 92] : Mayhew, D. J., *"Principles and Guidelines in Software User Interface Design"*, Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [Médini & al. 01] : Médini L., Chaillot M., Charlot J-M, *"Designing an Electronic Knowledge Book: How Advanced Knowledge Management Methods Can Help Information Modeling"*, Technical Communications, Special Issue on Information Modeling, sous presse (enfin !).
- [Mc Knight & al. 91] : Mc Knight, C., Dillon A., Richardson J., *"Hypertext in context"*, Cambridge University Press, Cambridge, 1991.
- [Motta & al. 00] : Motta E., Buckingham Shum S., Domingue J., *"Ontology-Driven Document Enrichment: Principles, Tools and Applications"*, in International Journal of Human-Computer Studies, vol. 52, n° 5, pp. 1071-1109, 2000.
- [Nanard & Nanard 91] : Nanard J., Nanard M., *"Using Structured Types to incorporate Knowledge in Hypertexts"*, in proceedings of the Hypertext'91 Conference, ACM Press, pp. 329-334, 1991.
- [Nanard & Nanard 93] : Nanard J., Nanard M., *"Should Anchors be Typed Too? An experiment with Macweb"*, in proceedings of the Hypertext'93 Conference, ACM Press, pp. 51-62, 1993.
- [Nanard & Nanard 98] : Nanard J., Nanard M., *"La conception d'hypermédias"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 15-34, 1998.
- [Nelson 67] : Nelson T. H., *"Getting it out of our system"*, in Information Retrieval: a Critical Review, G. Schechter (Ed.), Thompson Books, Washington D. C., 1967.
- [Nielsen 90] : Nielsen J., *"Hypertext & Hypermedia"*, Academic Press Inc., San Diego, Ca., 1990.
- [Normann 86] : *"User Centered System Design: New Perspectives on Human-Computer Interaction"*, Norman D-A., Draper, S-W. (Eds), Hillsdale, NJ: Lawrence Erlbaum Associates, 1986.
- [Nonaka 95] : Nonaka I., Takeuchi H., *"The knowledge-creating company"*, Oxford University Press, 1995.
- [Novak & Gowin 84] : Novak J., Gowin D. B., *"Learning how to learn"*, Cambridge University Press, 1984.
- [Palmquist & Kim 98] : Palmquist R. A., Kim K-S., *"Modeling the Users of Information Systems: Some Theories and Methods"*, Electronic Resources: Use and User Behavior, Hemalata Iyer (Ed.), The Haworth Press Inc., pp. 3-25, 1998.
- [Panijel 99] : Panijel C., Unité régionale de promotion et de formation pour l'information

- scientifique et technique de Paris / École des Chartes, Information Scientifique et Technique, disponible en ligne : <http://www.ccr.jussieu.fr/urfist/def-ist.htm>, 1999.
- [Peeters & Charlier 95] : Peeters H., Charlier P., *"Pour une sémio-pragmatique des hypertextes multimédia : proposition théorique de catégories d'analyse pertinentes"*, disponible en ligne : <http://www.comu.ucl.ac.be/GReMS/HugoWeb/semio-hypertext.html>, 1995.
- [Perlmann & al. 90] Perlmann G., Egan D. E., Relich K., Marchionini G., Nielsen J., Shneiderman B., *"Evaluating hypermedia systems"*, in proceedings of the ACM CHI'90 Conference on Human-Computer Interaction, Panel Session, pp. 387-390, 1990.
- [Pierce 78] : Pierce C. S., *"Collected Papers of Charles Sanders Pierce"*, 8 vol., Harvard University Press, Cambridge, Mass., 1978.
- [Prié 01] : Prié Y., *"Sur la piste de l'indexation conceptuelle de documents. Une approche par l'annotation"*, dans le numéro spécial de la revue Document Numérique sur l'indexation, sous la direction de J-M Jolion, Document Numérique, Vol 4, n° 1-2/2000, Hermès Science Publication, pp. 11-35, 2001.
- [Roger & al. 98] : Roger D., Lavandier J., Kolmayer E., *"Navigation et interfaces : cartes conceptuelles et autres outils"*, analyse bibliographique réalisée pour PARINFO (Programme d'Aide à la Recherche en INfOrmation), révision du 27 Avril 2000, disponible en ligne : <http://www.enssib.fr/bibliotheque/documents/travaux/kolmayer/kolm-navig0.html>.
- [Rouet & Tricot 95] : Rouet J-F., Tricot A., *"Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive"*, Sciences et techniques éducatives, vol. 2, n° 3, Hermès, Paris, 1995.
- [Rouet & Tricot 98] : Rouet J-F., Tricot A., *"Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 57-74, 1998.
- [Roussey & al. 99] : Roussey C., Calabretto S., Pinon J-M., *"État de l'art en indexation et recherche d'information"*, Document Numérique, Hermès, vol. 3, n° 3-4, pp. 121-149, Paris, 1999.
- [Quintana & al. 92] : Quintana Y., Kamel M., Lo A., *"Graph-Based Retrieval of Information in Hypertext systems"*, in proceedings of the Tenth International Conference on Systems Documentation, ACM Press, pp. 157-168, 1992.
- [Sakarovitch 84] : Sakarovitch M., *"Programmation discrète. Optimisation combinatoire, Méthodes mathématiques et algorithmiques"*, Hermann, Paris 1984.
- [Salton & Mc Gill 83] : Salton G., Mc Gill M-J. *"Introduction to modern information retrieval"*, McGraw-Hill, pp. 157-198, 1983.
- [Salminen & al. 95] : Salminen, A., Tague-Sutcliffe J., McClellan C., *"From Text to Hypertext by Indexing"*, in ACM Transactions on Information Systems, vol. 13, n° 1, pp. 69-99, janvier 1995.
- [Savoy 96] : Savoy J., *"Citation schemes in hypertext information retrieval"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 99-120, 1996.
- [Scapin 90] : Scapin D-L., Pierret-Golbreich C., *"Towards a method for task description: MAD"*, in L. Berlinguer and D. Berthelette (Eds), Work in display units 89, Elsevier

- Science Publishers, North Holland, 1990.
- [Shapira & al. 96] : Shapira B., Shoval P., Raveh A., Hanani U., *"Hypertext browsing: a new model for information filtering based on user profiles and data clustering"*, Online & CDROM Review, vol. 20, n° 1, 1996.
- [Shepherd 89] : Shepherd A., *"Analysis and training in information technology tasks"*, in Task Analysis for Human-Computer Interaction, Dan Diaper ed., Ellis Horwood Ltd., pp. 15-55, GB, 1989.
- [Shneiderman 76] : Shneiderman B., *"Exploratory experiments in programmer behavior"*, International Journal of Computer and Information Sciences, vol. 5, pp. 123-143, 1976.
- [Shneiderman 92] : Shneiderman B., *"Designing the User Interface: Strategies for Effective Human-Computer Interaction"*, 2^{ème} édition, Addison-Wesley, Reading, 1992.
- [Simoni & Fluhr 97] : Simoni J-L., Fluhr C., *"Accès à l'information à travers les graphes de termes"*, communication à ISKO-France, Lille, octobre 1997.
- [Simoni 00] : Simoni J-L., *"Accès à l'information à l'aide d'un graphe de termes construit automatiquement, intégration de l'interrogation et de la navigation"*, thèse de doctorat, Université Paris 7, Janvier 2000.
- [Smeaton 96] : Smeaton A. F., *"An overview of information retrieval"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 3-26, 1996.
- [Smith 96] : Smith W-J., *"ISO and ANSI Ergonomics Standards for Computer Products: A Guide to Implementation and Compliance"*, Prentice-Hall, 1996.
- [Taylor 62] : Taylor R., *"The process of asking questions"*, American documentation, pp. 391-397, 1962.
- [Tricot & al. 98] : Tricot A., Pierre-Demarcy C., El Boussarghini R., *"Définition d'aides en fonction des types d'apprentissages dans les environnements hypermédia"*, in Quatrième Colloque Hypermedias et Apprentissages, Poitiers, pp. 41-58, 15-17 octobre 1998.
- [Tricot & Nanard 98] : Tricot A., Nanard J., *"Un point sur la modélisation des tâches de recherche d'information dans le domaine des hypermédias"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 35-56, 1998.
- [Vignaux 94] : Vignaux G. : *"Les sciences cognitives : une introduction"*, Le Livre de Poche, La découverte, Paris, 1994.
- [Vogel 89] : Vogel. C., *"Génie cognitif"*, Collection Sciences Cognitives, Masson, Paris, 1989.
- [Von Bertalanffy 68] : Von Bertalanffy, L., *"General Systems Theory"*, Foundation, Development, Applications, New York: G. Braziller (1968), Trad. française: *"Théorie générale des systèmes"*, Paris: Dunod, 1973.
- [Vora & Helander 97] : Vora P. R., Helander M. G., *"Hypertext and its Implications for the Internet"*, Handbook of Human-Computer Interaction, M. Helander, T. K. Landhauer, P. Prabhu eds., Elsevier Science B.V., pp.877-914, 1997.
- [Vossen & al. 89] : Vossen P., Meijs W., Broeder M., *"Meaning and structure in dictionary definitions"*, Boguraev B., Briscoe T., éditeurs, Computational lexicography for Natural language Processing, pp. 171-190. Longman Group UK limited, London, 1999.

- [Waern & Hägglund 97] : Wærn, Y., Hägglund S., "*User Aspects Of Knowledge-Based Systems*", Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu (eds.), Elsevier Science, 1997, pp. 1159-1176.
- [Woodhead 90] : Woodhead Nigel, "*Hypertext & Hypermedia. Theory and Applications*", Sigma Press, GB, 1991.
- [Zarri & al. 99] : Zarri G. P., Bertino E., Black B., "*CONCERTO, An Environment for the "Intelligent" Indexing, Querying and Retrieval of Digital Documents*", in proceedings of the 11th International Symposium ISMIS'99, Fundation of Intelligent Systems, LNAI, 1609, Varsovie, pp. 226-234, Juin 1999.
- [Zizi & Beaudoin-Lafon] : Zizi M., Beaudoin-Lafon M., "*Accessing Hyperdocuments through Interactive Dynamic Maps*", in proceedings of the ACM ECHT'94, European Conference on Hypermedia Technologies, pp. 126-135, septembre 1994.

Chapitre 3. Méthodes et outils : disciplines impliquées et positionnement opérationnel de la thèse par rapport à la problématique de l'accès à l'information

Dans ce chapitre, nous positionnons l'ensemble des outils, conceptuels et opérationnels, dont nous aurons besoin dans les chapitres suivants. Il s'articule en trois sous-parties : définitions, techniques et application. Comme son nom l'indique, la première permet de préciser le sens dans lequel nous employons certaines notions souvent utilisées dans cette thèse. La deuxième détaille, d'un point de vue opérationnel les outils méthodologiques et techniques sur lesquelles s'appuie le travail présenté. La dernière présente la partie applicative de ce travail de thèse, qui a servi de base à la réflexion méthodologique présentée ensuite.

3.1. Définitions

Dans cette thèse, il sera fréquemment fait référence à des notions comme Information Scientifique et Technique, Système d'Information ou Gestion des Connaissances. Ces notions sont parfois complexes ou possèdent plusieurs acceptions. C'est pourquoi nous proposons ici une série de définitions. Celles-ci nous permettront de déterminer, parmi les différents points de vue existant dans la littérature, celui qui nous paraît le plus pertinent pour chaque notion par rapport au problème qui nous occupe.

3.1.1.1. Information

Pour les précurseurs de la théorie de l'information, comme [Shannon 48], la notion d'information correspond aux flux physiques qui sont véhiculés lors de la communication entre un émetteur et un récepteur. Ceux-ci peuvent être des humains en train de discuter, auquel cas, l'information circule entre eux sous forme d'onde sonore dont la manifestation physique est une variation de la pression d'air. Ils peuvent aussi être des ordinateurs qui communiquent via un réseau. Dans ce cas, l'information est transmise sous forme de bits, qui transitent grâce à la propagation d'une onde électromagnétique dans les câbles ou à travers l'atmosphère.

Cependant, si cette description physique de l'information est une base de travail utile dans des domaines comme l'électronique ou le traitement du signal, elle devient vite insuffisante lorsque l'on s'intéresse au contenu sémantique de cette information ou à l'utilisation qui en est faite par les individus. Dans cette optique, Gary Marchionini ([Marchionini 95]) définit l'information comme « anything that can change a person's knowledge ». Cette interaction entre information et connaissance rappelle alors les processus d'émission et d'interprétation de l'information décrits au chapitre précédent.

Plus précisément, [Buckland 91] distingue les concepts *information-as-process* (l'acte de communication, tel qu'il est défini par Shannon), *information-as-knowledge* (le contenu de l'information véhiculée lors de la communication, et qui correspond à la définition de Marchionini) et *information-as-thing* (le contenant de l'information qui permet de la stocker ou d'y accéder). Cette distinction rejoint celle qui est couramment faite en Français entre données, information et connaissance. [Nonaka 95] propose les termes de flux et de données pour caractériser respectivement les première et dernière facettes de l'information.

Ces deux concepts sont utiles pour la conception d'un système d'accès à l'information à des niveaux d'abstraction différents. Les flux dynamiques d'information qui circulent entre les différentes entités du système sont à prendre en considération lors de la phase de mise au point technique du système, tandis que l'ensemble des données stockées sont partie intégrante de l'architecture du système. C'est pourquoi, dans ce document, qui s'intéresse aux aspects méthodologiques de la démarche de conception, le terme d'information fera référence à un quelconque élément des données stockées. Lorsque nous nous intéresserons aux aspects dynamiques du système, nous emploierons les termes de flux d'information ou de transfert de données.

3.1.1.2. Information Scientifique et Technique

Nous restreindrons notre étude au champ de l'*Information Scientifique et Technique (IST)*. Selon la Direction de l'information scientifique, des technologies nouvelles et des bibliothèques ([DISTNB 98]) du Ministère de l'Enseignement Supérieur et de la Recherche, l'IST se définit comme « *l'information utile aux individus et aux organisations dans leur activité professionnelle. Même si le terme est moins usité au plan international, on peut également la qualifier, dans cette optique, d'information "professionnelle". Les champs de l'information spécialisée se distinguent donc en principe de l'information pour le grand public et de l'information à finalité pédagogique. L'acception "scientifique et technique" s'entend*

traditionnellement au sens large et recouvre aussi l'information technico-économique, technologique, juridique, réglementaire, les normes, les brevets, etc. » L'IST est donc considérée dans le cadre d'une activité professionnelle au sein d'une organisation, et autour d'un *domaine* particulier.

3.1.1.3. Systèmes d'information

De nombreux types d'informations peuvent être rencontrés en IST. En effet, pour Claire Panijel ([Panijel 99]), de l'Unité Régionale de Promotion et de Formation pour l'Information Scientifique et Technique (URFIST) de Paris, « *l'IST s'inscrit dans des formes documentaires diverses : revues scientifiques, thèses, rapports, actes de congrès, ouvrages spécialisés, manuels, bibliographies, résumés, prépublications mais aussi brevets, cartes, banques d'images, données statistiques et économiques... Elle circule également de manière plus informelle, souvent hors publication, dans les conférences, les échanges inter-personnels, le courrier électronique, les groupes de discussion, les sites web, sous forme de notes de travail, données provisoires, tableaux, graphiques, descriptions méthodologiques, programmes, appels d'offre etc.* ». L'ensemble de ces éléments, en rapport avec le domaine d'activité d'une organisation, est alors regroupé sous la dénomination de Système d'Information (SI).

Ici, le terme « système » est donc à considérer dans son acception ensembliste ou systémique ([Von Bertalanffy 1968], [de Rosnay 75], [Le Moigne, 77]) et non informatique (rencontrée par exemple sur les sites des entreprises qui commercialisent du matériel ou des logiciels). L'ensemble des outils permettant la production, le stockage et l'accès à l'information ne peuvent cependant être exclus de ce SI. Ils en constituent l'aspect dynamique. Un exemple de description du SI d'une organisation est proposé lors de la description du projet opérationnel présenté à la fin de ce chapitre.

3.1.1.4. Éléments du SI et granularité de l'information

Un élément du système d'information d'une organisation est donc un élément de donnée stocké sur un support et supposé accessible informatiquement ou physiquement. Ce peut donc être une entité telle qu'un livre, un fichier, une image. Mais cela peut aussi être une sous-partie de ces entités : un chapitre, un paragraphe, une ligne de code, un "zoom", voire le regroupement de plusieurs entités, comme une collection d'images ou un recueil de notes. En fonction des situations, une entité informationnelle peut donc être considérée comme un sur-ensemble d'informations, une information ou un sous-ensemble d'information. Le niveau de

granularité des éléments d'un SI n'est donc pas fixe et doit être spécifié pour en permettre l'accès.

3.1.1.5. Connaissance

Un individu doit accomplir un certain nombre de tâches qui composent son travail. Pour cela, il met en œuvre des *connaissances*. Ces connaissances font référence à un *modèle mental* ou une *représentation* du domaine qui lui sont propres. De nombreux travaux existent et ont donné lieu à de nombreuses définitions en sciences cognitives sur ces trois concepts ([Kosslyn 80], [Johnson-Laird 83], [Normann 88], [Finke 89]). Nous ne tentons pas ici de résumer ces études sur la complexité du fonctionnement de l'esprit humain. Nous n'avons besoin dans cette thèse que de savoir qu'elles constituent un système complexe (au sens de Le Moigne, c'est-à-dire comprenant de nombreuses relations entre ses éléments).

Cependant, le système de connaissances d'un individu ne nous est pas connu ; nous escomptons juste qu'il est fonction de l'ensemble des informations qu'il a préalablement reçues de son environnement. Ces informations touchent des domaines différents (individuel, culturel, linguistique ou environnement de travail). Dans cette thèse, nous faisons l'approximation que les connaissances liées à chacun de ces domaines sont indépendantes. C'est pourquoi seul le domaine de l'environnement de travail de l'individu nous importe ici. Nous désignons également cet environnement de travail par le terme de contexte d'émission (ou d'interprétation) d'une information.

Un second postulat que nous faisons ici est de considérer le processus d'émission d'information comme réversible. De ce fait, une information qui est émise par un individu (l'émetteur) pour formaliser une certaine connaissance⁴⁷ – si tant est qu'il existe un élément unitaire de connaissance – produit chez l'individu qui la reçoit (l'interprétant) une connaissance similaire⁴⁸. Par extension, nous admettons qu'il est alors possible de dupliquer une partie du système de connaissances d'un individu à un autre par la transmission d'un ensemble d'informations appropriées.

Le problème est alors de trouver cet ensemble d'informations. En effet, le fait de véhiculer l'information correspondant à la sous-partie du système de connaissances à transmettre peut

⁴⁷ Littéralement, il conviendrait d'appeler ce processus de transcription de la connaissance en information « informatisation ». Pour éviter la confusion avec l'acception courante de ce terme, inventée en 1962 pour traduire le terme anglais *computerization*, nous lui avons préféré le terme de formalisation.

⁴⁸ Toutes choses égales par ailleurs, c'est-à-dire, à condition que les systèmes de connaissances des deux individus soient similaires, à la connaissance formalisée près.

ne pas suffire à l'interprétant. En particulier, si le système de connaissances de ce dernier ne contient pas les connaissances auxquelles cette sous-partie doit être reliée, il ne peut être compris correctement par l'interprétant. Lors de la formalisation de l'information, l'émetteur fait implicitement le choix de la quantité de connaissance à transmettre pour que son message puisse être correctement interprété. Il arrive bien entendu que ce choix soit erroné et que l'information transmise ne suffise pas à l'interprétant pour se constituer le même schéma de connaissances ou produise des connaissances redondantes ou interférentes avec d'autres éléments de son système de connaissances.

3.1.1.6. Gestion des Connaissances

Dans ces conditions, nous définissons ici la Gestion des Connaissances comme une discipline qui s'applique à identifier l'ensemble des connaissances qui doivent être transcrites sous forme d'information pour générer chez l'interprétant la même sous-partie du système de connaissances. Cette discipline se fonde sur l'idée que les relations entre les éléments de connaissance correspondent aux interactions entre les différentes notions qu'elles symbolisent. Par conséquent, l'intégration d'une connaissance dans un système se fait si toutes les notions auxquelles est reliée la notion correspondant à cette connaissance font déjà partie du système de connaissances de l'interprétant.

Pour l'individu, l'ensemble des autres notions avec lesquelles chacune interagit constitue l'environnement de celle-ci. Dans le cadre d'une activité professionnelle, où ces notions sont manipulées les unes en fonction des autres, on parle de contexte d'utilisation des termes ou des concepts du domaine.

La Gestion des Connaissances a donc pour but de véhiculer, en même temps que l'information relative à une connaissance, l'information relative à son contexte d'utilisation, du point de vue de l'émetteur de cette information. Pour cela, elle doit préalablement disposer de l'ensemble des relations entre les concepts du domaine. La cartographie de ces relations est l'objet de la capitalisation (ou modélisation) des connaissances.

Dès lors, nous pouvons nous placer d'un point de vue extérieur et ne considérer que les flux d'information qui circulent entre les individus. Nous considérons donc l'individu comme une « boîte noire » qui émet ou reçoit de l'information. La Gestion des Connaissances vise alors à aider l'interprétation de cette information en la resituant dans son contexte d'utilisation.

En définitive, la Gestion des Connaissances n'intervient pas à un niveau cognitif mais à un niveau informationnel. Comme tout autre méthode de transmission de l'information, elle peut

alors s'accommoder des outils de traitement de l'information (et notamment les systèmes de Gestion Électronique de Documents) existants pour recontextualiser de l'information dans un certain contexte d'émission.

3.1.1.7. Ontologies et types de classification

D'après la définition du dictionnaire ([Robert 91]), l'ontologie en tant que discipline est la « partie de la métaphysique, qui s'applique à l'être en tant qu'être, indépendamment de ses déterminations particulières. » Plus précisément, pour les terminologues de l'approche viennoise, il s'agit de la discipline qui s'intéresse à un objet en fonction des relations intrinsèques qui le définissent ([Wüster 76], via [Felber 87], via [Van Campenhoudt 96]).

Une ontologie est l'application de cette discipline à un objet (ou domaine) particulier. Elle permet de définir cet objet par intension, grâce à des relations établies en fonction des propriétés jugées importantes dans un modèle associé à l'objet. Du point de vue des éléments liés par ces relations, une ontologie est alors une classification des composantes de l'objet. Les relations de classification ontologiques peuvent être de type partitives (fait partie de), de composition (est composé par), de contiguïté spatiale ou temporelle (se trouve près de, suit). Ces relations s'opposent aux relations logiques, qui sont définies par [Wüster 76] comme des relations de « ressemblance ». D'après la compréhension que nous en avons, Cette dénomination semble dépasser le concept d'égalité et englober celui d'inférence.

Une taxinomie (ou taxonomie) au sens de [Vogel 88] (p. 102) est un type de classification statique. Cette définition permet quasiment de l'inscrire dans le champ d'application de l'ontologie, si ce n'est que pour Claude Vogel, la taxinomie admet également une définition partitive d'un objet par extension. Pour cet auteur, la taxinomie est l'élément de base du savoir.

En pratique, le mode de classification employé dans les réseaux sémantiques distingue les liens ontologiques relatifs à la décomposition en sous-concepts ou à l'agrégation en entités génériques (ces liens sont dits verticaux) de toutes les autres propriétés pouvant induire des relations ontologiques (ces liens sont diagonaux). Les concepts de même niveau d'un graphe partagent les propriétés du sur-ensemble et sont alors liés par des relations logiques (ou horizontales) qui ne sont pas représentées.

La gestion des connaissances peut être considérée comme un sur-ensemble des ontologies, puisqu'elle incorpore des relations horizontales d'inférence qui dépassent la définition de ce concept. L'amalgame fait parfois à tort entre gestion des connaissances et ontologie est à

notre sens l'une des principales sources de réduction de cette première discipline à une classification statique des éléments du domaine.

Une actinomie est une définition intrinsèque d'un objet par des relations d'inférence dynamiques. Pour [Vogel 88], « l'actinomie est à l'action ce que la taxinomie est à la constatation. » Le recours à des actinomies permet une définition du fonctionnement synchronique interne d'un objet. L'actinomie est notamment un élément important de la décomposition systémique d'un domaine, visant à rendre compte de sa structure statique et de son fonctionnement dynamique.

3.1.1.8. Utilisateur

Dans cette thèse, nous considérons l'utilisateur comme l'un des individus qui seront amenés à utiliser le système d'accès à l'information scientifique et technique que nous voulons concevoir. Parmi les multiples autres dénominations possibles – sujet pour les psychologues, usager pour les documentalistes ou interprétant pour les sémiologues – nous avons choisi le terme d'utilisateur, en accord avec l'orientation informatique et ergonomique de cette thèse.

Un tel individu évolue dans un cadre professionnel à dominante scientifique et technique. Les conséquences de ce postulat sont les suivantes : sa fonction au sein de l'organisation est bien identifiée, il est soumis à une obligation de résultats dans son travail et à des contraintes de temps variables.

3.2. Disciplines

Dans ce paragraphe, nous abordons plusieurs disciplines d'un point de vue opérationnel. Il s'agit de positionner, pour chacune d'elles, les techniques développées qui nous ont servi pendant cette thèse sans que nous n'y apportions aucune modification. Dans les cas où plusieurs techniques existent et où nous en avons choisi une, nous justifions ces choix.

3.2.1. La Gestion des Connaissances

3.2.1.1. La méthode de gestion des connaissances MKSM

3.2.1.1.1. Les différents types de méthodes de gestion des connaissances

Au chapitre précédent, nous avons vu que la gestion de connaissances peut présenter une approche utile pour l'accès à l'information : en fournissant une description formalisée du contexte d'émission de cette information, elle aide l'individu qui la reçoit à l'interpréter en

fonction de ce contexte. Dès lors, même si nous avons souligné le fait que toutes les méthodes servent à gérer l'information et non la connaissance, différentes approches existent, et ont donné lieu à plusieurs méthodes de gestion des connaissances. Le tableau ci-dessous propose un classement de ces disciplines et une liste non exhaustive des méthodes de gestion des connaissances qui en sont dérivées.

Discipline	Méthodes de Gestion des Connaissances
Intelligence Artificielle, Sciences Cognitives	KADS [Wielinga & al. 93], [Schreiber & Wielinga 93], KOD, Rex, SAGACE, MKSM
Génie Logiciel	Merise [Tardieu 83 & al.], HOOD [Booch 92], SADT [IGL 82]
Ingénierie Documentaire, Gestion Électronique de Documents (GED)	AVIS ([Lippold & al. 94]), « bases de connaissances » en ligne (par exemple : sites web de support technique) ou en Intranet (bases d'informations partagées via Outlook®, Lotus Notes®)
Paradigmes opérationnels divers	<ul style="list-style-type: none"> - mémoire d'entreprise [Pomian 96] - analyse fonctionnelle [Sector 98] - analyse du vocabulaire technique [Kristen 98] traitement linguistique de la documentation technique ([Millerat & al. 96])

Tableau 2. Différentes méthodes de Gestion des Connaissances classées par leurs disciplines d'origine.

Toutes ces méthodes possèdent différentes caractéristiques : la qualité des fondements théoriques sur lesquels elles reposent, les méthodes de modélisation proprement dites, leur utilisabilité (i.e. facilité de mise en œuvre) et réutilisabilité dans des domaines différents, et leur opérationnabilité (i.e. utilisabilité au tant qu'élément dans une chaîne de conception visant à produire autre chose qu'un simple recueil de connaissances). Nous ne tenterons pas de dresser ici un classement de toutes ces méthodes en fonction de ces caractéristiques. Chacune d'entre elles a été développée dans un contexte particulier, et répond à un problème spécifique.

Parmi ces méthodes de gestion des connaissances, la méthode MKSM a été développée et est couramment utilisée à la DIST du CEA. Cette méthode bénéficie d'un retour d'expérience satisfaisant et d'un savoir-faire important dans cette unité. En outre, cette méthode s'appuie sur des formalismes diagrammatiques qui peuvent être mis au format hypermédia (voir plus loin) et donc permettre une stratégie d'accès à l'information par navigation. De plus, MKSM permet une modélisation aussi « large » que nécessaire de l'activité d'une organisation, ce qui nous sera utile pour le processus de conception présenté plus loin. C'est pourquoi nous avons

utilisé cette méthode de modélisation des connaissances comme base de travail pour cette étude de conception⁴⁹. Dans les paragraphes suivants, nous décrivons cette méthode.

3.2.1.1.2. *Les fondements théoriques de la méthode MKSM*

La méthode de Gestion des Connaissances MKSM, développée à la DIST du CEA, a été finalisée en 1996 par [Ermine & al. 96], mais s'appuie sur le retour d'expérience des projets menés à la DIST depuis 1990. Nous décrivons dans ce paragraphe les fondements théoriques de cette première version. Les paragraphes suivants exposent respectivement la version de la méthode utilisée dans cette thèse et les dernières évolutions de la méthode postérieure aux travaux présentés ici.

Dérivée de l'IA et des systèmes experts, la version originale de MKSM s'appuie sur plusieurs formalismes diagrammatiques et se fonde sur un outil conceptuel, imaginé en systémique par J. de Rosnay ([de Rosnay 75]) : le « macroscopie ». L'adaptation de cet outil à la Gestion des Connaissances est faite en considérant que la connaissance possède une dimension systémique⁵⁰ et une dimension sémiotique⁵¹. Ces deux dimensions sont chacune décomposées en trois points de vue qui forment un triangle.

La composante par laquelle J-L Ermine a choisi de commencer sa décomposition est la dimension sémiotique. Elle est composée des trois points de vue de Peirce : le signe, qui représente la trace physique de la connaissance (i.e. l'information), le signifiant, qui représente l'environnement en fonction duquel le signe considéré est interprété (le contexte) et le signifié, qui représente le sens associé par l'individu au signe physique (la connaissance). Cette décomposition est classiquement représentée dans un *triangle sémiotique* proposé dans la figure ci-dessous. Il est à noter que pour [Ermine & al. 96], l'étude du signe en tant qu'information relève du traitement de l'information. De nombreuses méthodes existent dans cette discipline, et le point de vue de l'information n'est donc pas traité par la méthode MKSM.

⁴⁹ Il est cependant probable qu'une méthode s'appuyant sur d'autres formalismes aurait pu conduire au même type d'étude. Dans le cadre de l'étude du statut des formalismes diagrammatiques spécifiquement aux problématiques d'accès à l'information et de modélisation du domaine, des modèles ne faisant pas partie de la méthode MKSM sont d'ailleurs considérés (cf. chapitre 5).

⁵⁰ D'où l'appellation de « systèmes de connaissances », utilisée dans la définition du terme connaissance. Cette dimension systémique met en avant l'importance des relations entre les différentes entités considérées, même si nous avons déjà souligné qu'il ne s'agit pas de connaissances mais d'informations.

⁵¹ Il existe ici une contradiction qu'il nous paraît important de souligner. Rappelons que la sémiotique est l'étude du signe, alors que la connaissance est par définition immatérielle, et donc décorrélée de l'information qui en est la manifestation physique. C'est l'une des principales raisons qui ont motivé les évolutions de la méthode MKSM présentées plus bas.

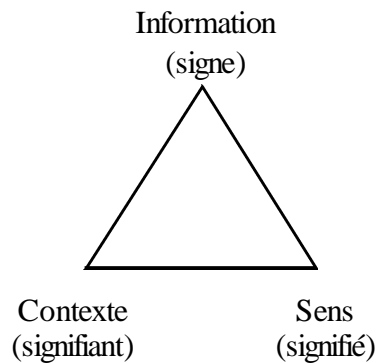


Figure 9. Le triangle sémiotique adapté à la méthode MKSM.

De la même façon, [Ermine & al. 96] proposent un *triangle systémique* qui décrit un élément d'un système en fonction des trois points de vue suivants : sa structure (son organisation d'un point de vue interne), sa fonction (par rapport aux autres éléments du système ou par rapport à l'extérieur) et son évolution (l'historique des différentes transformations qu'il a subies). Une des idées directrices sur lesquelles se fonde la systémique est que chaque élément d'un système se définit en fonction de ses relations avec les autres éléments. C'est pourquoi chacun de ces trois points de vue définit l'élément considéré en fonction d'un certain type de relations. La décomposition structurelle met l'accent sur les relations entre les composants internes de cet élément. Sa décomposition fonctionnelle fournit un modèle dynamique synchronique (i.e. à un instant donné) qui permet d'analyser le fonctionnement du système autour de cet élément. Son évolution est un modèle diachronique qui donne, élément par élément, un aperçu de l'évolution du système au cours du temps⁵².

Dans l'optique d'une définition de modèles diagrammatiques correspondant aux différents points de vue identifiés, cette décomposition permet déjà de définir les caractéristiques principales de ces modèles, en fonction de la composante systémique à laquelle ils se rattachent. Les modèles de structure sont des modèles statiques⁵³, tandis que les modèles

⁵² Pour être complet et respecter le paradigme systémique d'explicitation des relations entre les éléments, il faudrait pouvoir donner un modèle diachronique du fonctionnement synchronique du système, c'est-à-dire un modèle d'évolution du fonctionnement global du système. Un tel modèle permettrait de décrire la répercussion des évolutions de chaque élément sur ses relations avec les autres éléments et donc la répercussion de ces évolutions sur ces autres éléments. Un tel modèle serait assez complexe à mettre sous forme diagrammatique, et n'existe donc pas dans la méthode MKSM.

⁵³ Nous sommes conscient que cette remarque est en contradiction avec le caractère dynamique du modèle du domaine, présenté en 1996 comme structurant le contexte. Nous reviendrons sur ce point au chapitre 5.

synchroniques et diachroniques sont eux dynamiques. Cependant, aucun formalisme d'évolution diachronique n'est présenté dans la première version de la méthode MKSM⁵⁴.

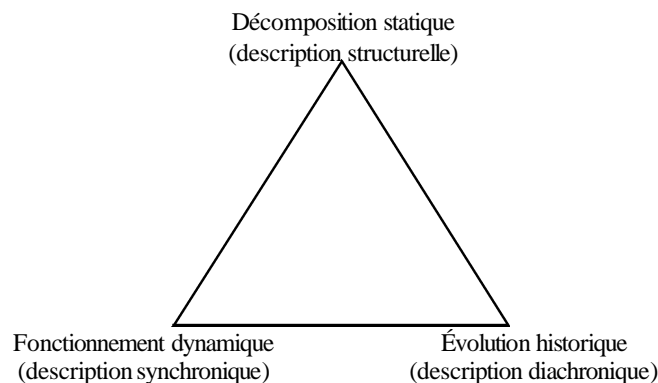


Figure 10. Le triangle systémique appliqué à la méthode MKSM.

En croisant ces deux dimensions, [Ermine & al. 96] obtiennent donc finalement quatre points de vue, correspondant aux structures et aux fonctions du contexte et du sens. Ces points de vue sont représentés dans la méthode MKSM par les modèles du domaine (pour la structure du contexte), des activités (pour la fonction du contexte), des concepts (pour la structure du sens) et des tâches (pour la fonction du sens). Le « macroscopie de la connaissance » est le référentiel sur lequel se fonde la méthode MKSM. Il repositionne ces modèles parmi les neuf points de vue originels et est présenté dans la Figure 11.

La méthode MKSM présentée par Jean-Louis Ermine et alii en 1996 propose également un autre outil conceptuel : le modèle Opérant Information Décision et Connaissance (OIDC). Ce modèle diagrammatique décrit l'articulation des deux principaux éléments hiérarchiques d'une organisation (l'opérant et le décisionnaire) en fonction des flux informationnels et cognitifs qui y circulent. Un modèle OIDC générique est présenté en Figure 12.

⁵⁴ J-L. Ermine admet l'utilité de ce point de vue, mais reconnaît ne pas avoir de formalisme diagrammatique adapté à proposer pour le mettre en forme. Ces formalismes ont été mis au point plus tard par [Van Craeynest 00].

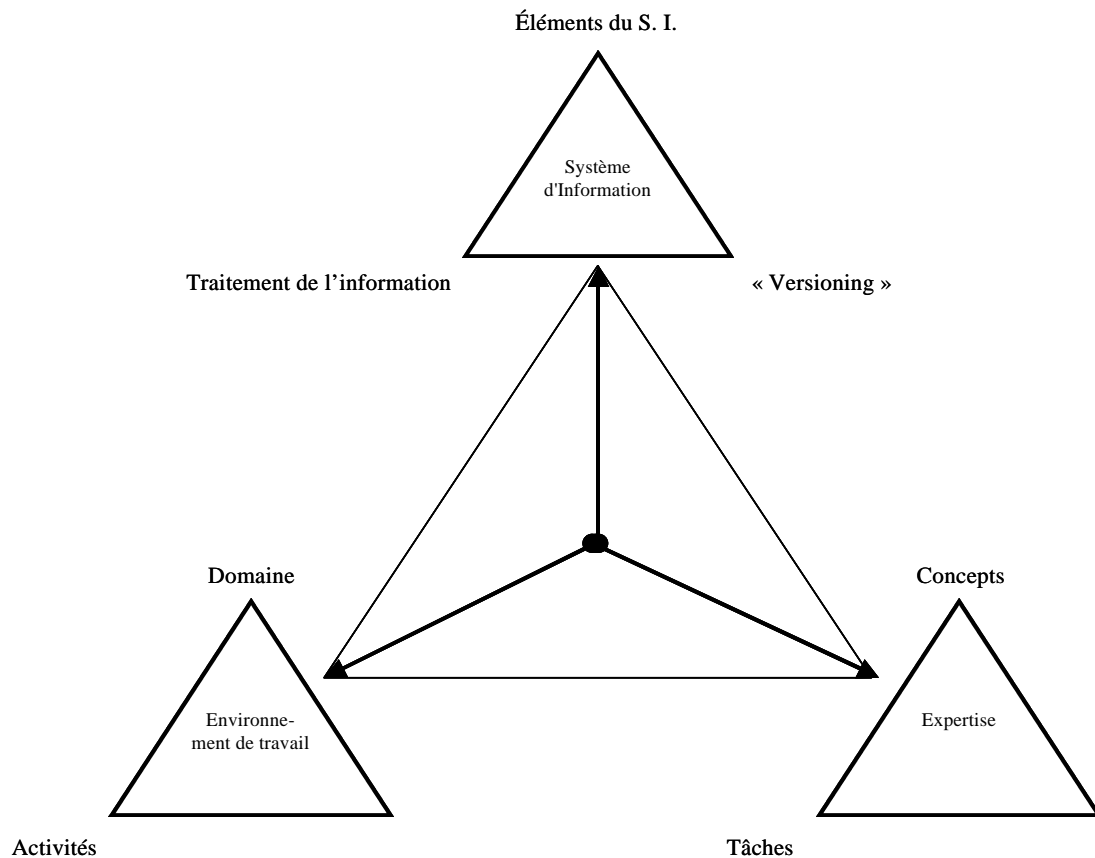


Figure 11. Le macroscopie de la connaissance présenté dans [Ermine & al. 96].

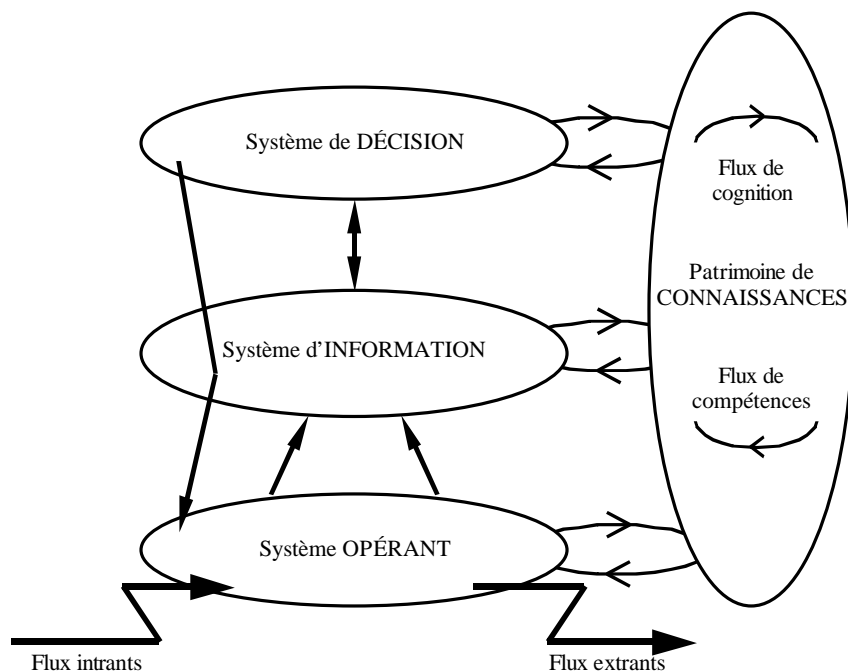


Figure 12. Le modèle OIIC de J-L. Ermine⁵⁵.

⁵⁵ D'après Jean-Louis Ermine, les flux cognitifs de ce modèle devraient englober les autres éléments du diagramme. Ce modèle se résume alors à une présentation de l'organigramme générique d'une l'organisation,

Ces deux derniers diagrammes sont donc des référentiels qui situent les modèles et les acteurs de l'organisation de connaissances les uns par rapport aux autres. Ils ont notamment été beaucoup utilisés par J-L. Ermine pour présenter la méthode aux décideurs des organisations internes et externes au CEA et « vendre » des projets de gestion des connaissances. Cependant, la méthode MKSM repose avant tout sur les formalismes diagrammatiques auxquels obéissent les diagrammes issus de la phase de modélisation. Ces formalismes sont présentés dans le paragraphe suivant.

3.2.1.1.3. *Les formalismes diagrammatiques de la méthode MKSM.*

Pour les chercheurs en Intelligence Artificielle, le diagramme est plus proche de la notion de connaissance que le langage, plus formalisé, qui se rapproche de celle d'information ([Glasgow & al. 95] (foreword), [Addis 96]). C'est pourquoi, en plus des paradigmes systémique et sémiotique, la méthode MKSM a recours à la diagrammatique pour favoriser la « compréhensibilité » de l'activité modélisée. Différentes façons de représenter l'information sous forme de diagrammes existent. La méthode MKSM présentée dans [Ermine & al. 96] retient quatre formalismes diagrammatiques pour représenter les points de vue décrits dans le paragraphe précédent.

3.2.1.1.3.1. **Le modèle du domaine**

Ce modèle a été créé pour décrire les phénomènes physiques⁵⁶ complexes autour desquels s'articule l'activité de l'organisation modélisée. La compréhension de ces phénomènes préside à celle de l'activité de l'unité. C'est à notre sens pour cela que ce modèle vient en tête dans le microscope MKSM.

Ce modèle se fonde sur le formalisme Source Cible Flux Champ (SCFC), développé spécifiquement pour la méthode MKSM pour mettre en évidence qualitativement la propagation des processus physiques. Chaque diagramme détaille un processus en évoquant un *système source*, relié à un *système cible* par un *flux* et en présence d'un *champ actif*. Ce modèle permet aussi d'énumérer les *événements déclencheurs* et *inhibiteurs* ainsi que les *conséquences* de ce processus. Un diagramme SCFC en sûreté-criticité est présenté ci-dessous.

qui est de peu d'intérêt dans la phase opérationnelle des projets. D'où le fait que nous n'avons pas eu recours à ce modèle.

⁵⁶ La plupart des projets ayant été menés au CEA dans des unités qui travaillent sur des sujets attenants à la physique nucléaire. D'autres projets traitent d'autres domaines scientifiques, comme la médecine ou la

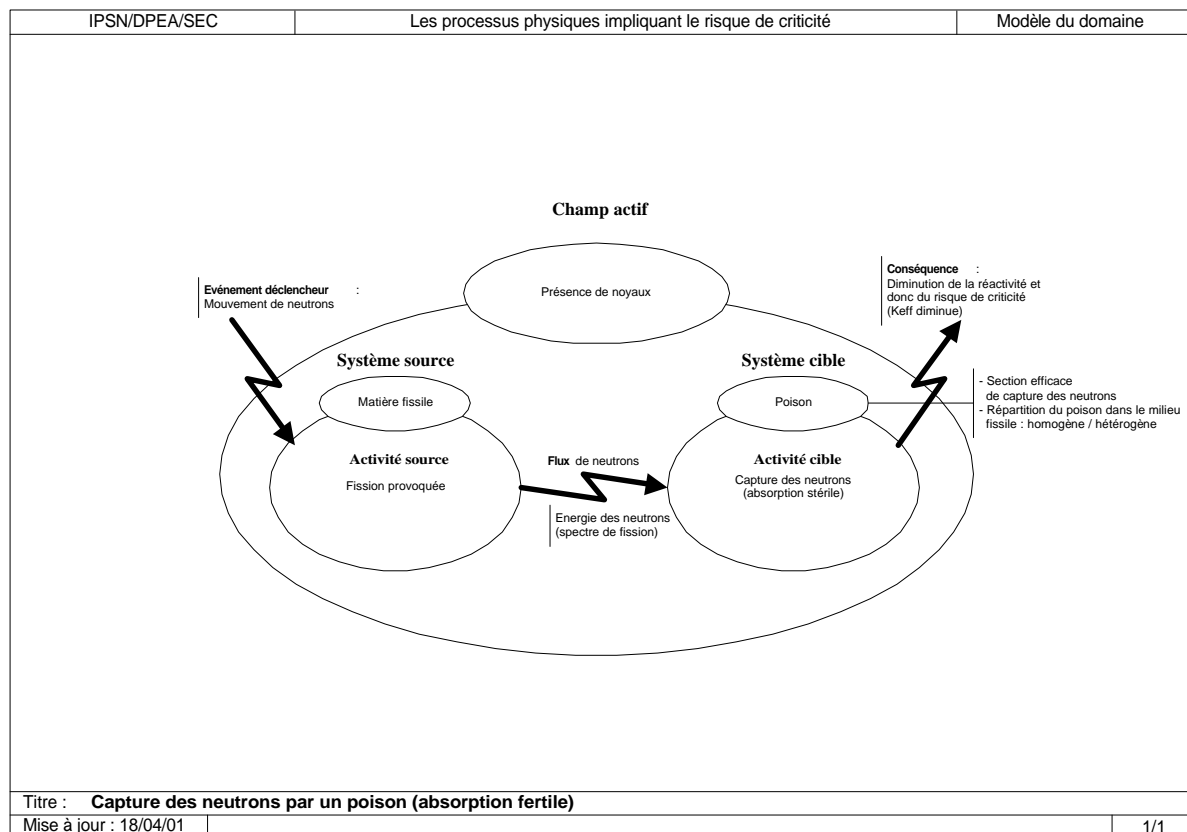


Figure 13. Diagramme SCFC de processus en criticité nucléaire. Dans ce type de diagramme, les types des champs sont explicitement nommés pour faciliter la compréhension du phénomène représenté par le lecteur.

Les événements déclencheurs et inhibiteurs ainsi que les conséquences d'un processus sont eux-mêmes d'autres processus qui sont décrits dans d'autres diagrammes. La méthode MKSM propose donc un autre formalisme, le *scénario des processus*, qui permet de positionner tous les processus les uns par rapport aux autres. Le scénario des processus physiques régissant la sûreté-criticité nucléaire est présenté dans la figure suivante.

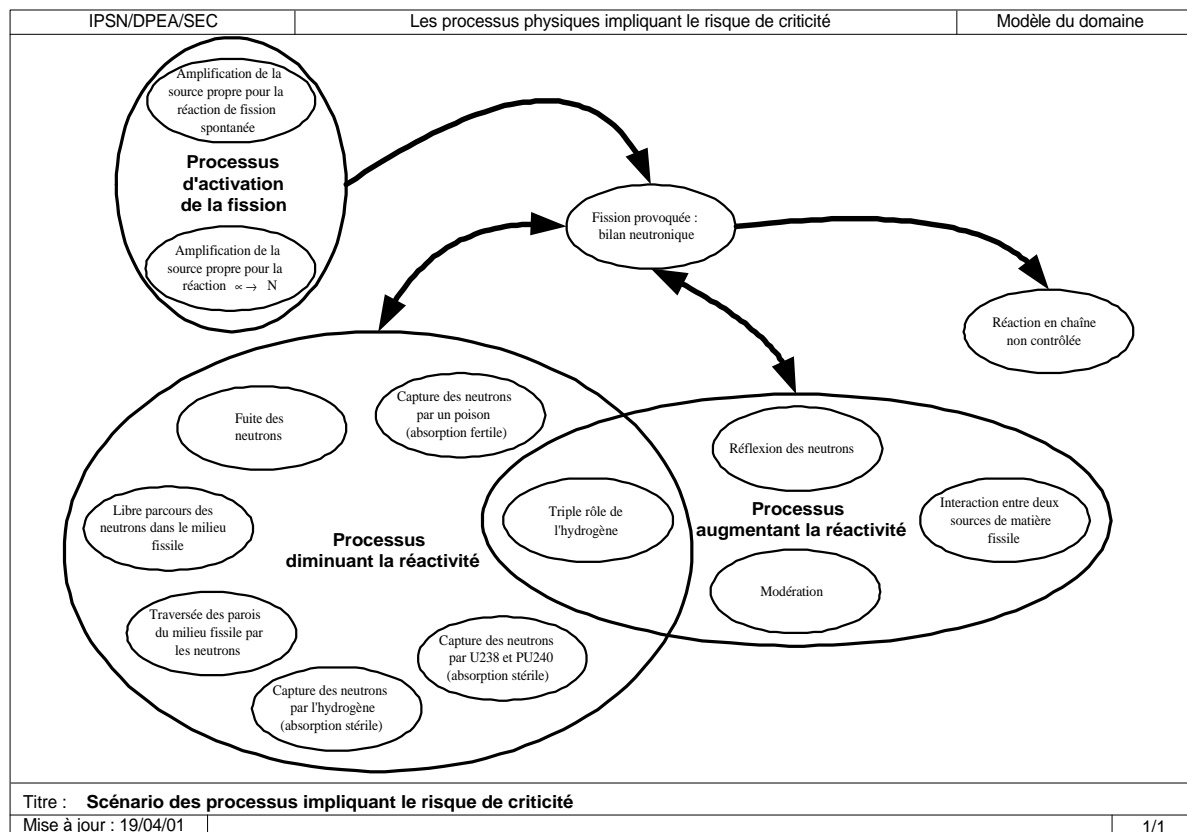


Figure 14. Scénario des processus impliquant le risque de criticité nucléaire. Ici, les flèches représentent des relations qualitatives de déclenchement potentiel d'un processus par un autre. Ce schéma est tiré du projet-pilote présenté plus loin. Il ne correspond toutefois pas exactement à la forme « classique » d'un scénario, telle qu'elle est définie dans [Ermine & al. 96].

Le modèle du domaine d'un système physique est donc composé d'un scénario et d'un ensemble de diagrammes SCFC détaillant tous les processus qu'il contient. Historiquement, ce modèle qualitatif des systèmes dynamiques découle des techniques utilisées en IA, notamment pour générer les règles des systèmes experts. Le caractère qualitatif de ce formalisme permet au lecteur de comprendre facilement le processus décrit. En revanche, il rend délicate son opérationnalisation dans les cas où des opérations quantitatives sont nécessaires (voir plus loin).

3.2.1.1.3.2. Le modèle des activités

Ce modèle est emprunté au formalisme SADT ([IGL 82]). Il permet de décrire chaque activité d'un acteur de l'organisation de deux façons : il la positionne dans un diagramme par rapport aux autres activités avec laquelle elle interagit et permet de la détailler dans un autre diagramme. Chaque diagramme présente en fait une activité en la décomposant en *sous-activités* et en mettant en évidence l'ensemble des *flux* et des *ressources*, matériels,

informationnels et cognitifs qu'elle met en jeu. Chaque sous-activité présentée étant décomposable, ce modèle est récursif et donc auto-structurant, puisque tous les diagrammes sont alors des décompositions à différents niveaux de profondeur d'une activité principale.

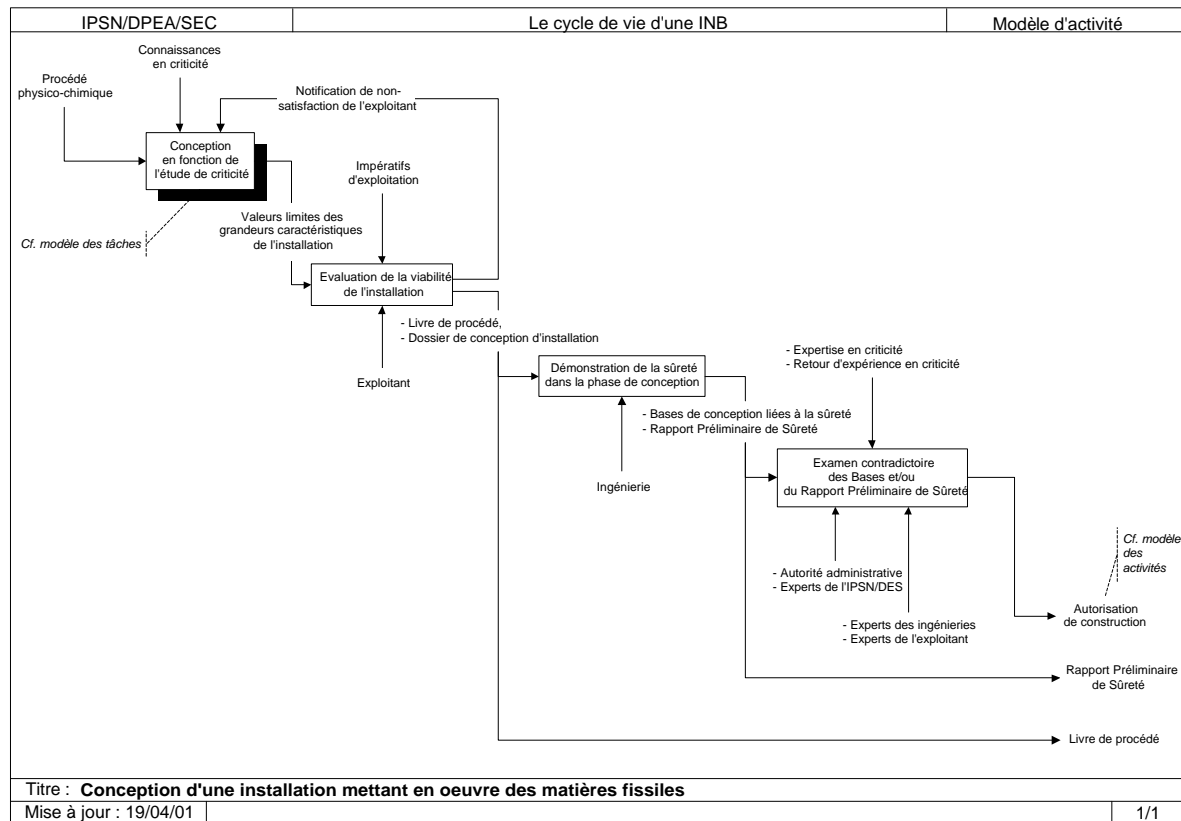


Figure 15. Diagramme d'activité pour les études de criticité pour la conception d'INB. Ce diagramme permet de positionner l'activité générique du criticien par rapport aux autres activités impliquées dans la conception d'une Installation Nucléaire de Base⁵⁷. Nous voyons par exemple sur ce diagramme que l'activité « évaluation de la viabilité de l'installation » est réalisée par l'exploitant nécessite la connaissance des « impératifs d'exploitation » et utilise les « valeurs limites des grandeurs caractéristiques de l'installation » pour produire un « livre de procédé » et un « dossier de conception d'installation ».

3.2.1.1.3.3. Le modèle des tâches

Ce modèle permet de décrire la stratégie cognitive d'un individu qui accomplit une tâche. Il est emprunté à la modélisation des tâches des utilisateurs pour la conception logicielle en ergonomie ([Scapin & Pierret-Goldbreich 90]). À la différence des activités, ce modèle ne

⁵⁷ Les activités sont symbolisées par les boîtes. Elles sont ombrées quand leur décomposition est donnée dans un autre diagramme. Les flux arrivant par la gauche d'une activité sont consommés par celle-ci tandis que ceux arrivant par le haut sont les ressources : ils sont nécessaires à l'exécution de l'activité sans toutefois y prendre part. Les flux entrant par le bas en sont les acteurs et ceux sortants (à droite), les produits.

présente pas les flux qui circulent entre les tâches⁵⁸. En revanche, il présente les *flots de contrôle* qui permettent à l'individu d'accomplir une séquence de tâches : des tâches peuvent être menées *séquentiellement*, en *parallèle*, de façon *itérative* ou *alternative* (i.e. à la suite du résultat d'un test). Comme le modèle d'activité, ce modèle est auto-structurant, puisque chaque tâche peut être décomposée en sous-tâches. Il peut d'ailleurs représenter un niveau de décomposition plus fin pour une activité et ainsi être cascadié avec le modèle des activités.

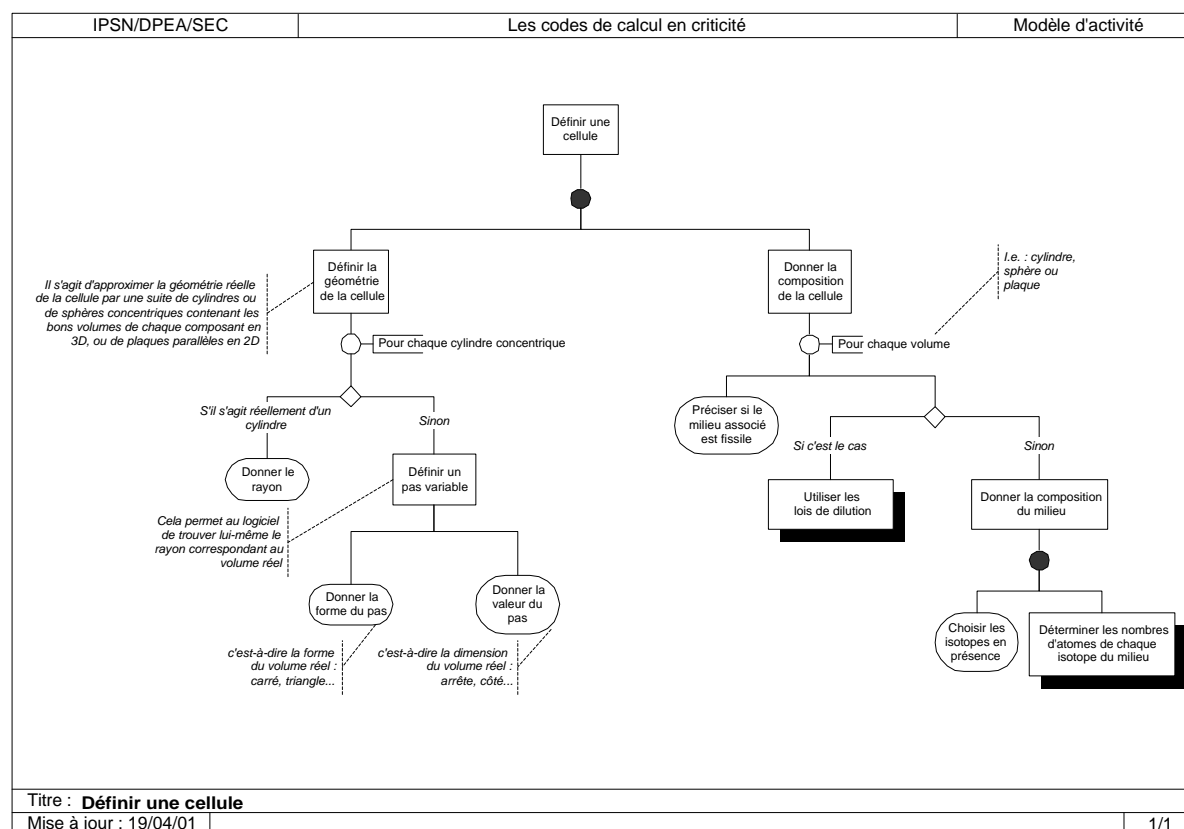


Figure 16. Diagramme de tâches pour les codes de calcul de criticité. Les flots de contrôle⁵⁹ permettent de détailler la stratégie cognitive de l'expert dans la réalisation de cette tâche. Par exemple, la tâche « donner la composition de la cellule » est réalisée de façon itérative en réalisant la tâche terminale « préciser si le milieu associé est fissile » et un test indiquant que la tâche à réaliser « si c'est le cas » est la tâche décomposée dans un autre diagramme « utiliser les lois de dilution ». « Sinon », il faut réaliser la tâche « donner la composition du milieu », qui se décompose elle-même séquentiellement en une tâche terminale et une tâche décomposée.

⁵⁸ Ces flux sont représentés dans le langage formel associé à la méthode MKSM, mais ont été retirés du formalisme graphique car ils nuisent à la compréhensibilité du graphe. Ce langage formel a été créé dans le but d'opérationnaliser la méthode MKSM. Il n'est plus utilisé, notamment car, comme cela a été démontré par B. Charreton, ce langage formel n'est que très difficilement opérationnalisable.

⁵⁹ Les boucles indiquent un processus itératif, les losanges une alternative, les disques noirs une suite de tâches à réaliser séquentiellement (par opposition au symbole // qui spécifie une parallélisation des tâches).

3.2.1.1.3.4. Le modèle des concepts

Ce modèle est issu de la recherche en sciences cognitives sur les structures sémantiques (cf. la définition des taxinomies, plus haut dans ce chapitre). Il présente une description arborescente d'un objet ou d'un concept, tel qu'il est vu par l'individu. Un concept peut comporter plusieurs *attributs* qui peuvent être nommés par *intension* ou énumérés par *extension*. Il peut également être décrit par des *spécialisations* ou des *généralisations* (ces deux relations sont inverses l'une par rapport à l'autre). Ce modèle permet également de décomposer chaque concept et est auto-structurant pour chaque concept de base décrit.

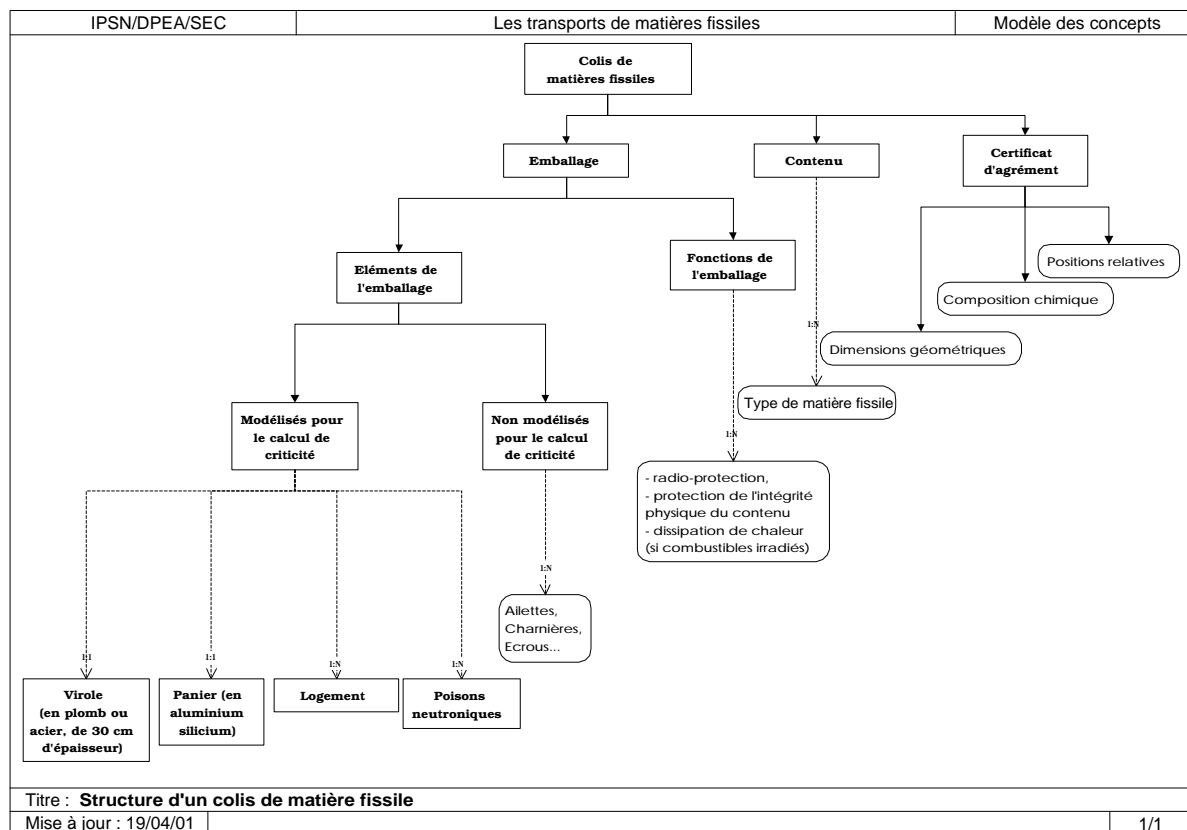


Figure 17. Diagramme de concepts pour les transports de matières fissiles. Ce diagramme présente le concept de colis de matière fissile vu par le critiqueur⁶⁰. Par exemple, le concept « éléments de l'emballage » se spécialise deux sous-concepts : les éléments « modélisés pour le calcul de criticité » et en éléments « non modélisés pour le calcul de criticité ». Le premier sous-concept est défini par intension (i.e. par une liste de ses caractéristiques) et le second par extension (i.e. par la liste des éléments qu'il définit).

Les paragraphes précédents présentent brièvement les modèles MKSM tels qu'ils ont été décrits dans [Ermine & al. 96], c'est-à-dire dans la perspective de la rédaction d'un « Livre de

⁶⁰ Les liens en traits pleins sont des relations de spécialisation. D'une manière générale, nous avons essayé de ne pas employer de généralisation, mais d'ordonner les concepts par ordre de décompositions croissantes. Les liens en pointillés représentent des liens de typage. Les concepts sont définis par intension dans les boîtes rectangulaires et par extension dans les boîtes aux coins arrondis.

Connaissances » papier⁶¹. Dans le chapitre 5, nous nous permettons d'émettre quelques considérations sur ces modèles et d'adapter leur organisation au problème qui nous occupe. Toutefois, dans le cadre du projet opérationnel décrit à la fin de ce chapitre, l'organisation de ces modèles n'est pas l'élément principal qui nous intéresse. En effet, le livre de connaissances produit ne fait pas état du macroscopique ; il se contente de présenter les différents points de vue en fonction des pôles d'activité modélisés.

Signalons de plus que ce livre de connaissances comporte un modèle supplémentaire, le modèle des objets, qui permet de donner une description physique des outils de travail auxquels l'expert fait référence et dont le formalisme est le même que celui des concepts. C'est la perception d'un manque de ce type de description qui nous a conduit à définir ce modèle. En effet, lors du projet-pilote sur la maîtrise du risque de criticité, nous avons eu besoin de matérialiser une distinction entre la description physique des objets et la perception conceptuelle qu'en ont les experts dans le cadre de leur activité. Par exemple, la matière fissile se définit physiquement par une certaine teneur en noyaux fissiles appartenant à une liste de radio-isotopes donnée. Pour les experts, elle est décrite par un mode de contrôle, un K_{eff} (voir, dans la partie concernant le projet-pilote, le paragraphe sur le domaine de la sûreté-criticité, plus loin dans ce chapitre) et un type de traitement à subir. Ce modèle a été élaboré en coopération avec les experts. Il correspond donc bien à un « trou » de la méthode. Le but de ce modèle est de permettre une meilleure accessibilité des notions utilisées dans le domaine aux individus non experts en « matérialisant » les descriptions conceptuelles données par les experts. Ce point est discuté en détail au chapitre 5

3.2.1.1.4. *Les évolutions de la méthode MKSM*

Depuis 1996, la méthode MKSM s'est enrichie de deux nouveaux formalismes, présentés dans [Van Craeynest & al. 00]. Il s'agit des modèles de l'historique des actions et de l'évolution des concepts. Ces modèles permettent de retracer une partie de l'historique de l'activité modélisée et des concepts attenants, en présentant des solutions qui ont été proposées puis adoptées ou abandonnées.

J-M. Van Craeynest définit donc ainsi deux des points de vue diachroniques que J-L. Ermine reconnaissait comme restant à établir lors de sa première définition de la méthode. Ces modèles n'ont pas été utilisés dans le projet-pilote présenté à la fin de ce chapitre, car leur

⁶¹ Voir plus loin.

mise au point est postérieure à ce projet. La figure suivante illustre l'ensemble des formalismes désormais disponibles depuis l'évolution 2000 de la méthode MKSM.

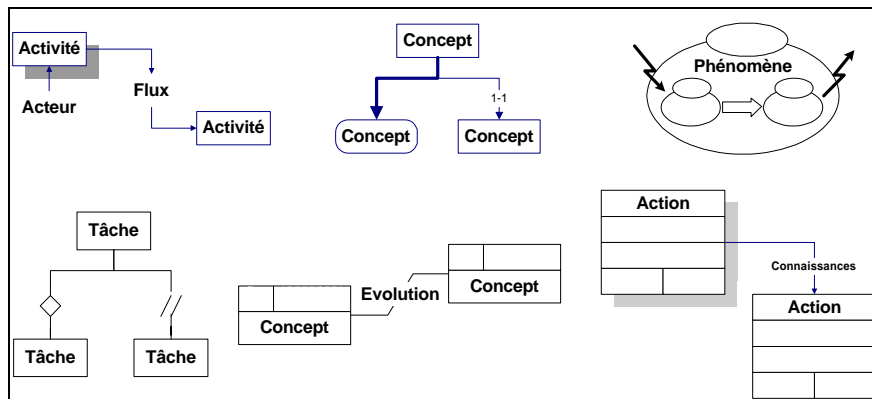


Figure 18. Formalismes utilisés dans la méthode MKSM : activités, concepts, phénomènes, tâches, évolution des concepts et historique des actions

3.2.1.1.5. *Méthodologie de modélisation*

En plus des formalismes diagrammatiques présentés ci-dessus, la méthode MKSM définit également une méthodologie de modélisation et de rédaction du livrable final : le Livre de Connaissances. Dans ce paragraphe, nous détaillons cette méthodologie en décrivant les techniques employées.

La phase de rédaction s'appuie sur des interviews d'experts du domaine. Le but de ces interviews est pour le chargé de projet d'obtenir de l'expert une description aussi détaillée que possible de son activité. Ces interviews sont aussi informelles que possible. Elles sont cependant enregistrées sur des cassettes audio, dans la mesure où l'expert donne son accord pour cela⁶².

Durant ces interviews, le chargé de projet interroge l'expert sur les différents aspects de son activité. Il guide la description de l'expert par le type de ses questions et structure ainsi le recueil de diagrammes. Il peut par exemple demander « pourquoi » il réalise une activité ou

⁶² Il peut arriver qu'un expert refuse de se laisser enregistrer. Un enregistrement sonore peut être perçu comme une trace dont il ne connaît pas l'utilisation ultérieure et qu'il ne veut pas voir diffusée. C'est alors au chargé de projet de l'assurer que ces cassettes ne sont utilisées que dans le cadre de la rédaction des diagrammes et qu'elles ne seront pas transmises à quelqu'un d'autre. Si l'expert refuse toujours de se laisser enregistrer, le chargé de projet n'a d'autre choix que de procéder à la rédaction des diagrammes en présence de l'expert, ce qui est beaucoup plus coûteux en temps pour ces deux individus. Cela est cependant assez rare, sans doute grâce au fait que le statut d'expert assure en général une certaine « tranquillité » vis-à-vis d'une telle diffusion de ces enregistrements.

Il est cependant fréquent que l'expert demande au chargé de projet d'arrêter l'enregistrement lorsqu'il a à expliquer quelque chose d'officieux qu'il ne veut pas voir figurer sur une cassette. S'il s'agit d'un élément

une tâche ; cela incite l'expert à la repositionner dans une activité ou une tâche plus globale. La question « comment », permet à l'expert de donner un niveau de décomposition plus fin d'une activité ou d'une tâche. Les mêmes types de questions existent pour orienter la rédaction de diagrammes suivant d'autres formalismes.

Une fois l'interview terminée, le chargé de projet utilise l'enregistrement pour rédiger ou corriger les diagrammes. Cette phase de rédaction correction est alors suivie d'une nouvelle interview, qui a pour but de valider ou d'infirmer les nouveaux diagrammes et d'obtenir des informations pour en rédiger de nouveaux. Cette alternance de phases d'interviews / rédaction se poursuit jusqu'à ce que l'expert et le chargé de projet estiment que l'activité est correctement modélisée et qu'un niveau de détail suffisant est atteint.

Ce processus itératif de rédaction / validation des diagrammes est finalement assez proche de la démarche de conception participative ([Von Hippel 88], [Mackay 88]) employée en conception logicielle pour évaluer la pertinence d'une interface ou d'une fonctionnalité vis-à-vis de l'utilisateur. Nous pouvons déjà noter que dans le cas où la modélisation des connaissances est une étape dans la conception d'un système d'accès à l'information dont les experts seront les premiers bénéficiaires, nous nous inscrivons pleinement dans ce type de démarche de conception.

3.2.1.1.6. *Utilisabilité et réutilisabilité de la méthode MKSM*

Le paragraphe précédent présente la méthodologie de mise au point d'un Livre de Connaissances. Nous nous intéressons dans ce paragraphe à l'intérêt que cette méthodologie présente pour les experts du domaine et les chargés de projet qui réalisent la modélisation en gestion des connaissances.

Du point de vue du chargé de projet, l'utilisabilité de cette méthode réside bien entendu dans la qualité du cadre de travail qu'elle fixe pour la modélisation. Chaque étape du processus de rédaction / validation est détaillée, et l'assimilation des techniques d'interviews, en permettant d'orienter l'expert sur le type de réponses à donner, permet de gagner énormément de temps à la fois dans l'explicitation de l'activité (interview) et dans sa mise en forme (rédaction).

Du point de vue de l'expert, l'intérêt de la méthode réside dans les formalismes diagrammatiques, qu'il arrive en général bien à s'approprier. Chacun des diagrammes synthétise un aspect précis de son activité, qu'il lui est aisé de valider ou d'infirmer. Dans ce

important de sa conception de son travail, le problème du chargé de projet est alors de trouver une façon acceptable d'intégrer cet élément dans les diagrammes.

dernier cas, il trouve facilement le point sur lequel il est en désaccord. La présence du chargé de projet aide également l'expert à expliciter son activité, ce qu'il est reconnu qu'il a du mal à faire tout seul. En effet, l'« angoisse de la page blanche » ajoutée à la méconnaissance du niveau de « non-expertise » du domaine des autres acteurs de l'organisation et, parfois, à une certaine réticence à se dessaisir de son savoir-faire, font en général des notes rédigées par un expert seul un outil très difficilement utilisable. De plus, le fait que les diagrammes sont rédigés par le chargé de projet et non par l'expert rend la modélisation moins coûteuse en temps de travail de l'expert, et pénalise moins l'activité de l'organisation.

Pour le chargé de projet, la réutilisabilité de la méthode est évidente. Les résultats empiriques et le retour d'expérience sont satisfaisants dans de nombreux domaines, essentiellement scientifiques et techniques. Les techniques de modélisation sont indépendantes du domaine. Elles peuvent cependant varier selon le type de livrable à produire, notamment dans le cas d'un système d'accès à l'information, comme nous le montrons plus loin. Dans ce cas, il ne s'agit plus d'un problème de réutilisabilité, mais d'opérationnabilité de la méthode MKSM.

3.2.1.1.7. Opérationnabilité de la méthode MKSM

La méthode MKSM est tout d'abord faite pour le transfert de connaissances entre les êtres humains. À ce titre, elle est essentiellement qualitative. De plus, la richesse des différents formalismes diagrammatiques se fonde sur les relations sémantiquement riches (mais de ce fait complexes) existant entre les éléments des diagrammes. Ces deux raisons la rendent difficilement utilisable pour un système informatique. Hormis pour les systèmes à base de règles, la nature calculatoire des logiciels usuels ne se satisfait pas de ce type de relations⁶³.

Des recherches ont cependant été menées dans ce sens, dont la plus aboutie semble être le « moteur de tâches » de B. Charreton ([Charreton & Ermine 96]). Cet outil utilise une grammaire BNF pour donner une transcription en langage objet des modèles MKSM. Mais même pour son auteur, le système élaboré est complexe et nécessite un complément quantitatif à la modélisation. De plus, il est centré sur une tâche bien précise : l'instrumentation du pilotage des systèmes de contrôle des domaines modélisés.

D. Malavieille a lui aussi utilisé la modélisation des connaissances pour la conception d'un système opérationnel. Dans ce cas, il s'agit de systèmes d'aide à la décision pour la gestion de

⁶³ Comme nous l'avons déjà souligné, c'est par exemple le cas pour l'opérationnalisation du modèle du domaine, où le système peut avoir besoin de comparer l'impact des effets déclencheurs et inhibiteurs de deux processus concurrents sur un troisième.

situations de crise, élaboré en collaboration avec le Service Départemental d'Incendie et de Secours de Seine Maritime ([Malavieille & al. 95]). Ce système se rapproche cependant des systèmes à base de règles, puisqu'il a pour but de renseigner un opérateur humain sur la conduite à tenir dans une situation donnée.

Mais à notre sens, la vraie problématique de l'opérationnalisation d'une méthode de gestion des connaissances repose dans sa finalité. Au départ, ces méthodes sont conçues pour capitaliser l'expertise des individus dans une organisation et la présenter sous une forme statique. Dès lors que la modélisation obtenue doit être réutilisée dans un autre but, nous voyons qu'il est nécessaire d'apporter des modifications (compléments ou suppression d'éléments) à la méthodologie de cette modélisation. L'un des objectifs de cette thèse est de définir les modifications d'ordre méthodologique à apporter à la méthode MKSM pour la rendre opérationnalisable dans le cadre de la conception d'un système d'accès à l'information.

3.2.1.2. Le Livre de Connaissances Électronique

3.2.1.2.1. *Le Livre de Connaissances*

La méthode de gestion des connaissances MKSM a pour but de produire un livrable appelé Livre de Connaissances (LC). Pour [Ermine & al. 96], il s'agit d'un recueil de toutes les connaissances du domaine qui ont été formalisées. À la lumière de la distinction que nous avons faite entre informations et connaissances au début de ce chapitre, nous préférons parler d'une explicitation de l'activité ou/et du contexte de travail des experts de l'organisation. Le LC est donc une entité informationnelle destinée à décrire aussi finement que possible l'activité modélisée.

Un LC est un document papier⁶⁴ qui contient l'ensemble des diagrammes rédigés, mais aussi plusieurs entités textuelles qui organisent ou complètent ces diagrammes : tables des matières, dictionnaires des termes spécifiques du domaine, fiches documentaires, glossaires, ainsi que des détails sur certains aspects de la modélisation que les experts ont voulu faire figurer sous forme textuelle plutôt que diagrammatique (ces derniers ont alors le même statut que les diagrammes). Outre tous ces éléments, un LC contient beaucoup de références d'un élément à un autre. Nous avons vu en effet que la compréhension de la complexité d'un domaine se fait également par l'assimilation des liens qui relient l'ensemble des concepts du domaine.

⁶⁴ Parfois assez volumineux : le LC réalisé par la DIST du CEA sur le procédé SILVA fait plus de mille pages.

L'ensemble de ces éléments et des liens qui les relient est alors supposé constituer une entité qui suffit à décrire le domaine modélisé. Un LC permet au lecteur d'appréhender l'activité décrite dans son ensemble et dans sa complexité. C'est alors un excellent outil de formation pour les personnels novices dans une organisation. Il permet aussi aux personnes extérieures au domaine de se faire une idée de l'activité d'une organisation. Il peut par exemple être utilisé comme « vitrine », pour présenter le domaine à des journalistes ou des personnes extérieures à l'organisation mais reliées à l'activité (fournisseurs, clients, hiérarchie).

Le niveau de détail avec lequel sont décrits certains aspects permet également de garder une trace du savoir-faire technique des experts. Le LC peut donc également être employé pour garder une trace de ce savoir-faire, lorsqu'un expert quitte l'organisation (par exemple dans le cas d'un départ à la retraite) ou lorsqu'un projet s'arrête, et que l'on craint de perdre le savoir-faire qui y a été développé.

3.2.1.2.2. *Le Livre de Connaissances Hypermédia*

Un Livre de Connaissances comporte donc un recueil de diagrammes décrivant l'activité des experts interviewés. La richesse des diagrammes et la compréhension des relations entre eux permettent au lecteur d'appréhender la complexité du sujet traité. Présenté sous format électronique, le Livre de Connaissances permet la navigation hypermédia entre les modèles de connaissances. L'utilisateur peut ainsi se construire une vue globale du domaine d'activité, centrée sur les aspects qui l'intéressent ou explorer un sous-domaine précis. Un tel outil, baptisé Livre de Connaissances Hypermédia (LCH), matérialise, sous forme de liens hypermédiés, les relations sémantiques définies dans la méthode MKSM. En parcourant ces liens, l'utilisateur peut « naviguer dans le domaine », et sa compréhension de celui-ci s'en trouve améliorée.

Pour [Ermine 97] (p.25), un LCH devient dès lors « un bon outil de présentation et de formation ». La transcription dans un format hypermédia d'un livre de connaissances permet une consultation plus aisée d'un contenu « dont le volume rend délicat la consultation linéaire. »

D'un point de vue hypermédia, le modèle correspondant à un LCH est classique. Nous reprenons ici la classification des concepts structurels d'un hypermédia, telle qu'elle est indiquée dans [Balpe & al. 96], p. 20 et présentée au chapitre 2. Aux éléments de base indiqués ci-dessous ont parfois été rajoutées des cartes conceptuelles pour faciliter la navigation dans les différents prototypes déjà réalisés.

Les nœuds du LMCE sont les diagrammes de connaissance MKSM ou des documents textuels particuliers destinés à préciser certains éléments des diagrammes. Chaque diagramme est rédigé selon l'un des formalismes de la méthode MKSM, définie dans [Ermine & al. 96]. Nous assimilons les documents textuels de l'hypermédia à des diagrammes rédigés selon un « formalisme textuel ». La structuration de ces nœuds est hiérarchique, et décompose le domaine d'activité de l'organisation en sous-domaines, formalismes et diagrammes de connaissance.

Les liens sont unidirectionnels et proviennent des relations sémantiques définies dans la méthode MKSM. Le typage de ces liens en fonction des différents types d'éléments des diagrammes sources et des types de diagrammes destinations est détaillé dans [Médini 97]. D'un point de vue générique, ceux-ci peuvent être typés en liens intra-diagramme, inter-diagrammes, inter-modèles, ou inter-sous-domaines.

Les ancres sont les éléments des diagrammes. Ces ancres sont naturellement typées en fonction des formalismes des diagrammes auxquels elles appartiennent. Remarque : ces ancres sont encore appelées « nœuds des diagrammes », dans la terminologie utilisée en gestion des connaissances.

Les premiers exemples de livres de connaissances hypermédiés ont été réalisés par [Le Huynh 95] avec le logiciel Toolbook® et par [Picard 96] avec Ilog Views®. Le récent essor des technologies W3 rend désormais les LCH (désormais réalisés au format HTML) diffusables sur un poste équipé de logiciels de bureautique standards. La portée de la modélisation étant l'organisation, le niveau de diffusabilité visé est alors l'intranet de cette organisation.

Du point de vue de l'environnement de génération de LCH, seuls les outils de mise au format hypermédia doivent être rajoutés à ceux de génération de LC papier. La plupart de ces outils, s'ils n'existent pas déjà (comme la fonction « exporter en HTML » du logiciel Visio®, utilisé pour rédiger les diagrammes), ne sont pas très délicats à mettre en place. Cependant, avec le nombre de projets menés à la DIST, un besoin récurrent d'homogénéiser tous ces outils s'est fait sentir. Il s'agit de mettre au point un atelier de génération de LC, englobant tout le processus de conception (de la rédaction à l'édition, au format papier ou HTML). Un tel atelier de génération de LCH (l'outil MNESIK®) a très récemment été mis au point par la

société C-Log, en collaboration avec la DIST du CEA⁶⁵. Nous reviendrons plus loin sur cet atelier de conception.

3.2.1.2.3. *Le Livre de Connaissances Électronique*

Relié au fonds documentaire de l'organisation modélisée, un LCH peut aussi être utilisé comme un système d'accès à l'information : en naviguant dans les diagrammes de connaissances, l'utilisateur passe en revue les éléments des diagrammes et peut choisir, à partir de chacun d'eux, soit d'utiliser des liens statiques qui aboutissent spécifiquement à certaines informations, soit d'interroger dynamiquement le fonds documentaire ou le système d'information s'y rapportant. Dans ce cas, chaque élément des diagrammes de connaissances permet d'accéder à une partie du système d'information de l'organisation.

Les principes fondateurs d'un outil d'accès à l'information reposant sur une modélisation de type MKSM, baptisé *Livre de Connaissances Électronique* (LCE), sont posés dans [Chaillot & Ermine 97]. Le Livre de connaissances présente le contexte d'utilisation de l'information (initialement, ces auteurs parlent du « sens associé à l'information », mais d'après les définitions au début de ce chapitre, nous considérons le terme de « contexte d'utilisation » comme plus approprié). En naviguant dans ce contexte, l'utilisateur se construit une vue globale du domaine d'activité et l'utilise comme une carte qui lui permet de se diriger vers l'information qu'il recherche. Pour cela, [Chaillot & Ermine 97] propose de coupler le document hypermédia réalisé avec un « système de recherche d'information ». Le LCE génère alors une requête à partir d'un nœud d'un diagramme et les réponses renvoyées correspondent aux documents pertinents.

La collection de diagrammes prise dans son ensemble permet alors de réaliser une indexation conceptuelle de la totalité de ce système d'information. La Figure 19 propose une description d'ensemble de l'activité de navigation et d'accès à l'information avec le LCE vue par l'utilisateur, revue par Mathias Chaillot en mai 2000.

⁶⁵ Du fait de la livraison tardive de cet atelier, nous n'avons pas pu en disposer pour ce travail de thèse. C'est l'une des raisons pour lesquelles le prototype présenté plus loin est l'une des premières versions développées et n'a pas été réactualisé depuis sa mise au point en septembre 99.

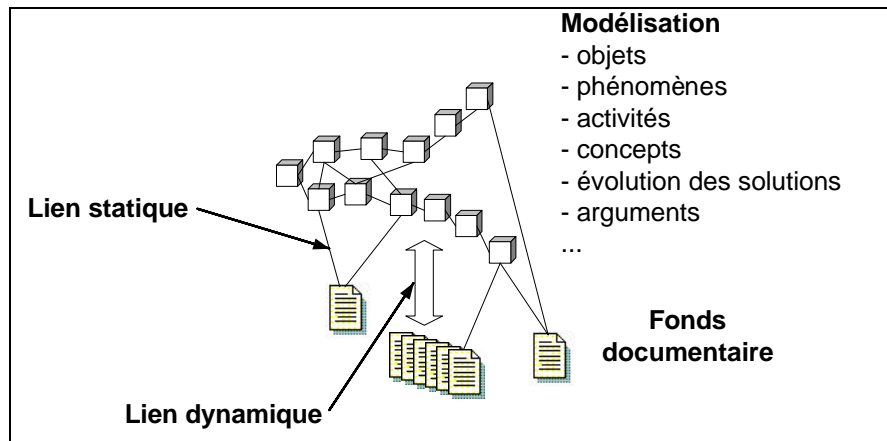


Figure 19. Liens statiques et dynamiques entre la modélisation multi-points-de-vue et les documents

Les liens statiques sont posés manuellement. Ils résultent d'une association sémantique, identifiée par les experts lors de la phase de modélisation, entre les éléments des modèles de connaissances et le contenu des documents. Ces liens sont a priori unidirectionnels, même si la fonctionnalité de retour au diagramme à l'origine d'un lien est prévue dans le LCE. Pour ces liens statiques comme pour l'ensemble des liens hypermédias entre les diagrammes, la sémantique associée est présentée à l'utilisateur avant leur activation. Cela lui permet de choisir la relation la plus pertinente par rapport à ses « buts de navigation » ([Chang & Rice 93]) ou à son problème d'information ([Marchionini 95]). Lorsqu'il utilise la fonctionnalité de retour, c'est à lui d'inverser la sémantique du lien précédemment emprunté.

Les liens dynamiques utilisent un moteur de recherche plein texte pour identifier soit les diagrammes pertinents par rapport à un élément textuel sélectionné par l'utilisateur dans un document, soit les documents pertinents par rapport aux contenus et aux types des éléments des diagrammes. Cette deuxième fonctionnalité du LCE, qui constitue son originalité en tant qu'outil d'accès à l'information, est développée au chapitre 4.

Nous avons participé à la réalisation d'un prototype de LCE à la DIST du CEA ([Charlot & al. 00]). Ce prototype a été présenté à la conférence internationale sur la criticité organisée par le SEC à Versailles ICNC'99 ([Médini & al. 99]). Il permet la navigation hypermédia dans les diagrammes de connaissances et l'interrogation d'une base documentaire ou l'envoi d'une requête sur le web (grâce au moteur de recherche AltaVista®), à partir d'un élément d'un diagramme. Ce prototype est constitué d'un ensemble de pages HTML dynamiques (i.e. utilisant des scripts JavaScript) préalablement générées avec l'atelier Arelice, développé en

Java par Jean-Marc Charlot. Nous présentons au chapitre suivant certaines de ses particularités et indiquons notre contribution à ce prototype⁶⁶.

Nous présentons dans cette thèse une version plus aboutie du LCE, baptisée Livre Multi-utilisateurs de Connaissances Électronique (LMCE), et en proposons une étude de conception. Cet outil se fonde entre autres sur l'hypothèse qu'un couplage fort diagrammes de connaissance / fonds documentaire améliore l'accès à l'information ainsi que sa compréhension. Cet outil propose également une prise en compte des caractéristiques des différents utilisateurs et la composition dynamique de l'interface en fonction de ces caractéristiques. Pour cela, il s'appuie sur une méthode de conception ergonomique spécifique et fait appel à la technique des documents virtuels personnalisables.

3.2.2. L'Ergonomie

La méthode de gestion des connaissances MKSM a été créée pour élaborer des livres de connaissances papier. Dans ce contexte, le format de présentation de ces livres de connaissances et les formalismes diagrammatiques utilisés sont appropriés pour ce type de livrable. Dans cette thèse, une fois la modélisation des connaissances effectuée, il s'agit de mettre au point un outil d'accès à l'information qui s'appuie sur cette modélisation. Dès lors, nous ne sommes plus simplement en présence d'une problématique de rédaction des diagrammes ou d'élaboration du livre de connaissances mais de conception logicielle.

Nous avons vu au chapitre 2 que pour ce type d'outil, une démarche de conception ergonomique est essentielle pour aider à la transmission de l'information. La démarche de conception présentée au chapitre suivant ne prend toutefois en compte l'utilisabilité de l'interface qu'à un niveau limité. Les aspects visuels de cette interface (critères de lisibilité de [Bastien & Scapin 93] ou [Bastien & al. 98]) ne sont pas discutés dans cette étude.

Deux composantes sont à prendre en compte pour cette démarche de conception⁶⁷ : d'une part, la définition d'une interface ergonomique pour l'accès à l'information à partir des diagrammes de connaissance et d'autre part, l'adaptation du système à différents types d'utilisateurs. Après avoir positionné l'ergonomie par rapport au processus de conception logicielle, nous abordons dans cette partie les différents aspects ergonomiques de la conception auxquels nous avons eu recours dans cette thèse.

⁶⁶ La description des travaux réalisés plus loin s'accompagne de l'identification de leurs auteurs. Le « nous » de narration employé ici désigne le rédacteur de cette thèse. Les autres auteurs sont explicitement nommés.

⁶⁷ Ces éléments seront développés au chapitre suivant.

3.2.2.1. Positionnement et historique de la discipline

Une définition de l'ergonomie est l'étude et la facilitation de l'utilisation (on parle dans cette discipline d'*utilisabilité*) de tous les types d'outils aux utilisateurs humains. En Anglais, cette définition est d'ailleurs très simplement résumée dans le nom de la discipline : « human factors », qui concernait à l'origine la conception des matériels destinés aux soldats américains pendant la seconde guerre mondiale. L'accent était alors mis sur l'étude des interactions physiques avec ces outils. Le terme « ergonomics » est lui plus usité dans le cadre de l'étude des problèmes d'utilisabilité des outils dans un environnement de travail. Cette discipline s'intéresse à l'utilisabilité de ces outils d'un point de vue physique, mais également psychologique (stress généré, aspects routiniers...). Notons que contrairement aux idées reçues, cette discipline est relativement ancienne, puisqu'il existe des traces de préoccupations d'ordre ergonomiques émises en 1828 par l'Institution of Civilian Engineering de Londres et que la première apparition du terme date de 1857.

L'ergonomie est donc a priori la discipline la plus directement concernée par les aspects humains de ces problèmes de compréhension et d'interprétation de l'information, dès lors que sa transmission s'effectue avec l'aide d'outils spécifiques. La branche de l'ergonomie qui s'occupe de favoriser l'utilisabilité des outils informatiques est l'ergonomie des logiciels, ce dont se préoccupe également l'Interaction Homme-Machine (IHM)⁶⁸.

On trouve sur le site de l'Association for Computer Machinery (ACM) la définition suivante de cette discipline : « (...) a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them. » ([ACM 92]) Cette définition est présentée comme pragmatique, dans la mesure où, pour ses auteurs, il n'existe pas de définition sur laquelle s'entendent tous les intervenants concernés par cette discipline.

Pour D-L. Scapin ([Scapin 97]), cette discipline date du début des années 1980 (voir [Scapin 86] pour une bibliographie des textes fondateurs de cette discipline). Elle se généralise cependant avec l'apparition d'interfaces graphiques sur les systèmes informatiques. Ces interfaces font suite aux interfaces de type « console », qui ne permettent aux utilisateurs que de taper du texte sur un écran noir et de visualiser les réponses du système sur ce même périphérique. Ces interfaces laissent assez peu de latitude à l'amélioration de l'interaction de l'utilisateur avec le système, et c'est pourquoi on situe l'origine de cette discipline à

⁶⁸ En Anglais, Human-Computer Interaction (HCI ou CHI).

l'apparition d'un type d'interfaces plus évolué. En fait, les premiers développements d'interfaces graphiques datent du début des années 1960, avec notamment le « sketchpad » de [Sutherland 63], destiné à permettre l'utilisation d'un stylo et d'un écran à tube cathodique. Cependant, ces développements n'ont pas alors pour but d'étudier l'interaction des utilisateurs avec de telles interfaces, mais uniquement de développer les techniques qui permettent ces interactions.

Pour l'ACM, l'ergonomie des logiciels fait appel à de nombreuses disciplines, tant du côté de l'informatique que des sciences humaines : « *On the machine side, techniques in computer graphics, operating systems, programming languages and development environments are relevant. On the human side, communication theory, graphic and industrial design disciplines, linguistics, social sciences, cognitive psychology, and human performance are relevant. And of course, engineering and design methods are relevant* » ([ACM 92]). L'ergonomie des logiciels est devenue une composante à part entière de la conception informatique, en abordant la problématique de conception du point de vue de l'utilisateur et non plus uniquement des traitements.

Nous avons vu au chapitre 2 les différents apports de l'ergonomie des logiciels pour la conception informatique. Dans les paragraphes qui suivent, nous présentons la dimension ergonomique de notre approche de conception du LMCE.

3.2.2.2. La méthode de conception ergonomique appliquée à la conception du LMCE

Un livre multi-utilisateurs de connaissances électronique est un outil combinant les stratégies d'accès à l'information par interrogation et par parcours. Concevoir un tel outil passe donc à la fois par la conception d'un hypermédia et par celle d'un système de recherche d'information.

3.2.2.2.1. Conception du Livre de Connaissances Hypermédia

Initialement (i.e. pour [Bush 45] ou [Nelson 67]), un hypertexte est une collection d'unités informatives élémentaires, entre lesquelles sont définis des liens. La mise au point d'un hypertexte réside donc principalement dans la définition de ces éléments informatifs et de ces liens.

L'information signifiante de la partie hypertextuelle d'un LMCE réside dans les diagrammes de connaissances. Le contenu de ceux-ci est élaboré selon la méthode MKSM, qui est le point de départ de cet outil. Il n'est donc pas dans nos objectifs d'y apporter de modification.

La méthode MKSM ainsi que la littérature s'y rapportant fixent également le contenu sémantique des relations existant entre les différents formalismes diagrammatiques (voir par exemple [Médini 97]). Cependant, leur mise en œuvre sous forme de liens hypermédias est laissée à l'appréciation du concepteur. Il est toutefois précisé que l'explicitation de la sémantique de ces liens est nécessaire pour permettre à l'utilisateur de comprendre l'articulation des schémas et donc se construire une vision globale du domaine représenté.

La liste des méthodes ergonomiques de conception d'un hypermédia est donnée dans l'état de l'art (chapitre 2). Le contenu informationnel et relationnel de la partie hypermédia d'un LMCE étant fixé par avance, nous avons choisi, pour nous aider à les mettre sous ce format, de suivre des règles de conception existantes. La littérature propose de nombreuses règles ergonomiques pour la conception d'hypermédias. En particulier, pour les hypermédias dédiés au web, [Bastien & al. 98] citent [Fucella & al. 98], [Fucella & Pizzolato 98] et [Vora 98]. La démarche de conception présentée au chapitre 4 suit les critères ergonomiques proposés par [Bastien & al. 98]. Ces auteurs s'inspirent de [Leulier & al. 98] et présentent des recommandations de conception et des méthodes de validation, adaptés aux hypermédias simples et également aux outils plus évolués (i.e. dynamiques).

3.2.2.2. Conception du système d'accès à l'information

Outre la collection de diagrammes de connaissances mise sous forme hypermédia, le LMCE comporte également un système d'accès à l'information complexe, dont les fonctionnalités sont définies et étudiées au chapitre 4. Contrairement aux diagrammes de connaissances, nous ne connaissons pas à l'avance ces fonctionnalités, ni les attentes des futurs utilisateurs. C'est pourquoi nous ne pouvons utiliser de méthode de définition de l'interface a priori. Il nous faut :

- modéliser l'activité des utilisateurs du futur système,
- identifier leurs attentes par rapport à ce dernier,
- déterminer les fonctionnalités pertinentes pour un outil d'accès à l'information fondé sur une modélisation de l'activité,
- définir l'interface adaptée à chaque situation,
- évaluer l'utilisabilité de cette interface et au besoin la modifier, et

- évaluer l'utilité du prototype et conclure sur la pertinence de l'approche.

Pour schématiser, nous pouvons décrire la chaîne de spécification, de conception et de validation du LMCE par le diagramme d'activité ci-dessous. Les paragraphes suivants détaillent chacune des étapes de ce processus.

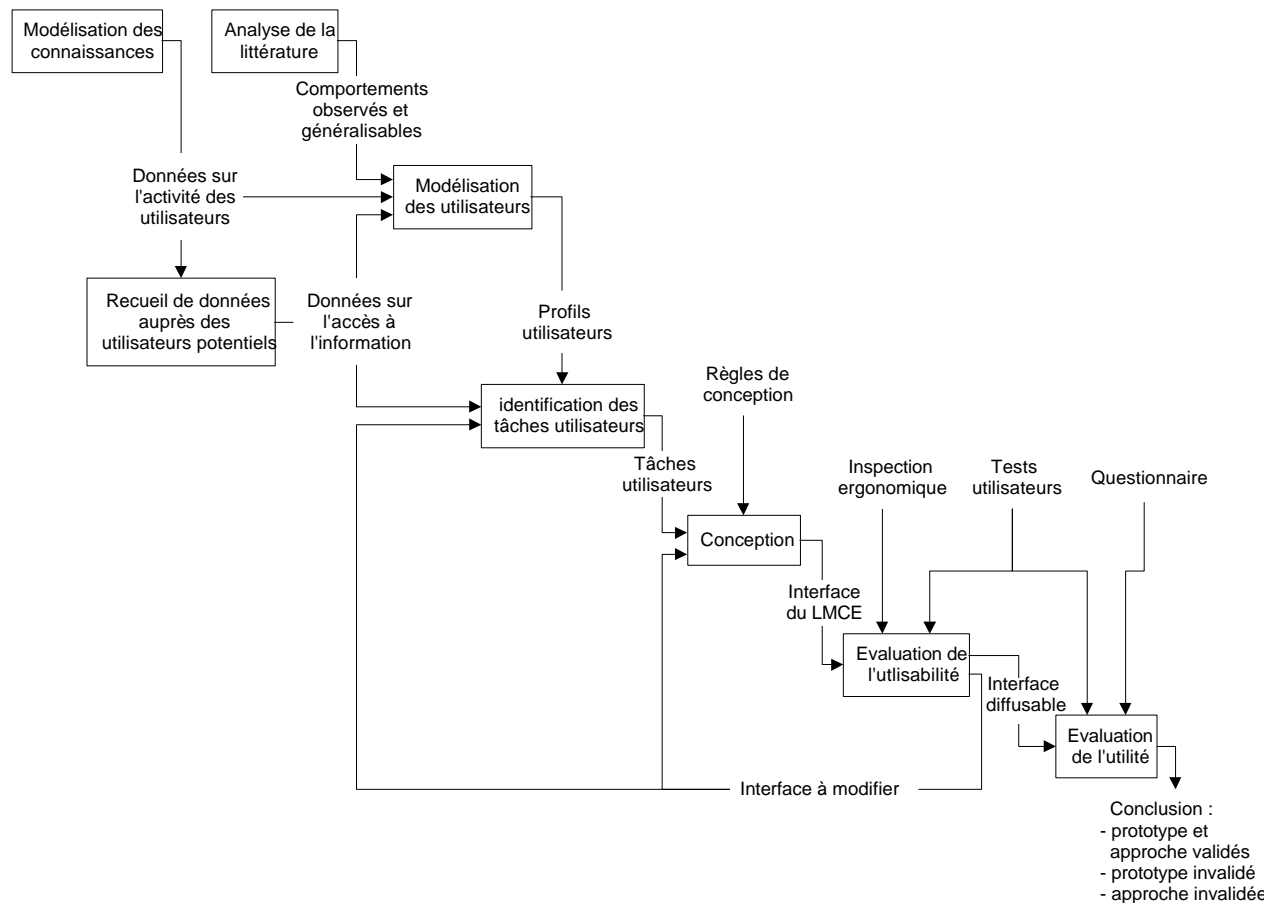


Figure 20. Schéma général de conception de l'interface du LMCE⁶⁹.

3.2.2.2.2.1. Utilisation de la gestion des connaissances pour une conception justifiée de l'interface

L'objectif de ce paragraphe est de présenter la méthode de conception utilisée pour la mise au point des aspects structurels et fonctionnels de l'interface du LMCE. La conception de cette interface s'appuie sur une modélisation des caractéristiques et des besoins des utilisateurs (voir ci-dessous). Nous montrons au chapitre suivant comment cette modélisation peut être incluse dans la phase de modélisation des connaissances avec la méthode MKSM. Nous

⁶⁹ Nous n'avons pas représenté ici les aspects fonctionnels de cette conception, comme l'ensemble des diagrammes issus de la modélisation pour ne pas surcharger la figure. Il est évident qu'ils sont nécessaires à l'activité de conception de l'interface.

disposons alors des données nécessaires à la conception sous forme de diagrammes MKSM. Il nous semble donc intéressant de poursuivre l'étude de conception à l'aide des mêmes formalismes. Une justification de cette approche est donnée a priori dans la problématique de la thèse (cf. chapitre 4). Une discussion de la pertinence de cette démarche de conception est proposée au chapitre 5.

Ces diagrammes permettent de présenter les différentes étapes de la conception. Ils ne permettent cependant pas de justifier les choix de conception, ce qui nous apparaît indispensable dans le cadre de la démarche de conception ergonomique du LMCE. Une telle justification permet d'éviter certaines erreurs de conception en explicitant les raisonnements qui mènent aux différents choix effectués.

L'idée de la justification des choix de conception est à la base de l'approche du Design Rationale⁷⁰ ([Lee & Lai 91], [Buckingham Shum 96], [Fisher & al. 96], et plus généralement, [Moran & Carroll 96]). Pour cela, plusieurs formalismes ont été mis au point dans cette discipline, comme les formalismes IBIS ([Kuntz & Rittel 70] via [Fisher & al. 96]) ou PHI ([Mc Call 79] via [Fisher & al. 96]). Ces formalismes sont reconnus comme complexes et ne nous sont pas apparus comme utilisables dans le cadre de cette démarche de conception pluridisciplinaire.

Nous avons en outre tenté de justifier ces choix à l'aide du formalisme QOC (pour Questions, Options and Criteria, cf. [Mc Lean & al. 91], [Farenc & Palanque 99]). L'utilisation de ce formalisme ne nous dispensait cependant pas d'un complément textuel à cette justification. C'est pourquoi nous avons fini par abandonner l'idée d'une justification diagrammatique des choix de conception.

La méthode de conception présentée ici s'appuie donc sur les formalismes diagrammatiques MKSM pour identifier les différentes étapes de la conception. Les choix permettant d'aboutir à ces étapes sont explicités textuellement.

3.2.2.2.2. Modélisation des besoins en information

La méthode de modélisation employée se fonde sur un recueil et une analyse préalables des données concernant les tâches d'accès à l'information avant l'utilisation du LMCE. La méthode de recueil des données s'appuie sur la modélisation des connaissances pour établir la

⁷⁰ L'idée de base du Design Rationale est d'assurer a priori la pérennité et a posteriori la traçabilité des choix de conception en explorant la totalité de l'*espace de conception* (design space) relatif à un *problème* (issue) donné.

modélisation ergonomique des tâches des utilisateurs du système. Nous décrivons cette méthode plus en détails dans les chapitres suivants.

3.2.2.2.2.3. Modélisation des utilisateurs

L'une des améliorations notables apportées dans ce travail à la définition du LCE donnée dans [Chaillot & Ermine 97] est la prise en compte des spécificités des utilisateurs. En effet, plusieurs utilisations différentes d'une modélisation de l'activité d'une organisation présentée au format hypermédia et couplée à un moteur de recherche peuvent être envisagées (cf. chapitre suivant). De plus, le projet-pilote présenté plus loin modélise l'activité de différents experts, dans différents sous-domaines d'activité de l'organisation modélisée. Les besoins en informations des différents utilisateurs varient donc en fonction du sous-domaine. Le LMCE réalisé doit donc satisfaire des utilisateurs aux préoccupations différentes.

Pour [Waern & Hägglund 97], l'adaptation de l'information présentée aux caractéristiques (« psychological requirements ») de l'utilisateur nécessite la définition d'un modèle de celui-ci (« user model »). Du fait de la multiplicité des caractéristiques des utilisateurs potentiels du LMCE, nous définissons des *profils d'utilisateurs* qui vont permettre à l'interface de réagir spécifiquement en fonction des attentes et des besoins de l'individu avec lequel elle interagit. Trois types de caractéristiques apparaissent comme pertinentes pour la définition de ces profils ([Wahlster & Kobsa 89], [Woodhead 90] (p.136), [Kass & Finin 88], [Lainé-Cruzel 94], [Waern & Hägglund 97]).

Le niveau d'expertise de l'utilisateur en fonction des différents sous-domaine de l'activité modélisée : nous avons défini arbitrairement trois niveaux d'expertise (expert, apprenant et novice) pour chaque utilisateur dans chaque sous-domaine, en accord avec les experts des domaines modélisés. Les différents sous-domaines sont déterminés par la modélisation des connaissances.

Les préférences de l'utilisateur en matière de recherche d'information : différentes stratégies d'accès à l'information ainsi que différents degrés de filtrage de l'information peuvent être préférés par les utilisateurs ([Allen 97]). Pour [Foucault 99], ces préférences ne sont pas nécessairement induites par les niveaux d'expertise de ce utilisateurs. Cette composante est donc une dimension indépendante de la définition des profils des utilisateurs.

Les préférences ergonomiques et de visualisation de l'utilisateur : ces préférences sont indépendantes du domaine modélisé. Nous nous référons à la littérature pour définir les

Cette exploration est obtenue par un passage en revue aussi exhaustif que possible des solutions envisageables et

caractéristiques des utilisateurs qui semblent être les plus déterminantes sur l'utilisabilité d'un système d'accès à l'information ([Woodhead 90], [Mc Knight & al. 91] (p. 55), [Allen 97]). Dans ce travail, nous avons choisi de nous limiter à la mise en page, au type et au nombre d'informations de natures différentes présentées en même temps à l'écran pour chacune des différentes tâches réalisées avec le LMCE⁷¹. Les choix d'attribution de ces préférences aux différents types d'utilisateurs sont discutés au chapitre 4. À terme, l'ambition est de disposer d'un « éditeur de préférences de visualisation » qui permette à chaque utilisateur de spécifier des paramètres personnalisés pour chacun des éléments d'interface qu'il rencontre. D'autres préférences de visualisations pourront alors être prises en compte (couleurs, polices de caractères...) Un tel éditeur n'a pas été réalisé dans ce travail.

3.2.2.2.2.4. Identification des tâches utilisateurs

L'identification des tâches utilisateurs est réalisée grâce au modèle des tâches de la méthode MKSM, lui-même inspiré de MAD. Les pré et post conditions peuvent cependant être indiquées par des diagrammes flot de données (i.e. des diagrammes d'activités). Cette identification des tâches s'appuie sur la modélisation des connaissances, selon une approche décrite au chapitre suivant.

3.2.2.2.2.5. Conception de l'interface

La méthode de définition de l'interface à partir des spécifications de tâches identifiées dans l'étape précédente se fonde sur des règles de conception. Comme pour le LCH, cette partie de l'application est au format HTML (dynamique). C'est pourquoi nous utilisons la même série de règles que précédemment (i.e. celles citées dans [Bastien & al. 98]).

3.2.2.2.2.6. Évaluation et validation du prototype et de l'approche

La phase d'évaluation et de validation du prototype fait appel à une inspection ergonomique, à des tests utilisateurs et à un questionnaire. Cela n'ayant pas été réalisé dans ce travail, le détail du déroulement de cette phase est donné dans les perspectives à court terme, au chapitre 5. Comme cela est mentionné au paragraphe 2.3.4.4, le choix de segmentation de l'évaluation en plusieurs méthodes a pour but de limiter le coût de cette phase de la conception.

par l'argumentation de chacune d'elles.

⁷¹ D'autres types de préférences sont sans doute intéressantes à prendre en compte, et en particulier, nous aurions souhaité nous intéresser aux différents modes de navigation hypermédia (cf. [Claverie 97]), mais nous n'avons ni les compétences pour cela, ni le temps de les acquérir.

3.3. Les technologies employées pour la réalisation du prototype

3.3.1. La technologie W3 et XML

Cette technologie est bâtie sur la famille des « markup languages » (langages de balises) : Standard Generalized Markup Language (SGML), HyperText Markup Language (HTML) et eXtensible Markup Language (XML). Les spécifications, normes et recommandations concernant ces langages sont disponibles sur le site du W3C⁷². L'intérêt de ces langages est principalement leur portabilité. C'est ce qui a guidé le choix de la DIST pour ce type de formats de sortie du LCE.

Plus précisément, nous nous sommes orienté vers le format XML non seulement pour la modularité des structures de données qu'il autorise, mais également pour la possibilité qu'il offre, en conjonction avec le langage de feuilles de style associé (XSL), de permettre la génération de documents HTML aux contenus différents à partir de la même structure de données. Nous indiquons plus loin l'intérêt du couplage d'un document XML (et de sa DTD), avec des feuilles de style (XSL et les feuilles de style « cascades » CSS) pour la réalisation de documents virtuels personnalisables qui permettent la personnalisation de la présentation des différents éléments d'interface en fonction du profil de l'utilisateur.

3.3.2. L'approche DVP

Les Documents Virtuels Personnalisables (DVP) sont des assemblages formatés de documents ou de parties de documents, d'après un modèle donné. L'aspect virtuel de ces DVP vient du fait que ces documents n'existent pas sur un support physique, mais sont générés à la volée. Les DVP peuvent alors être définis en fonction des besoins en information ou des préférences générales d'un utilisateur. Cette approche est fondée sur la réutilisabilité des données et des structures de données et permet de séparer physiquement la forme du contenu.

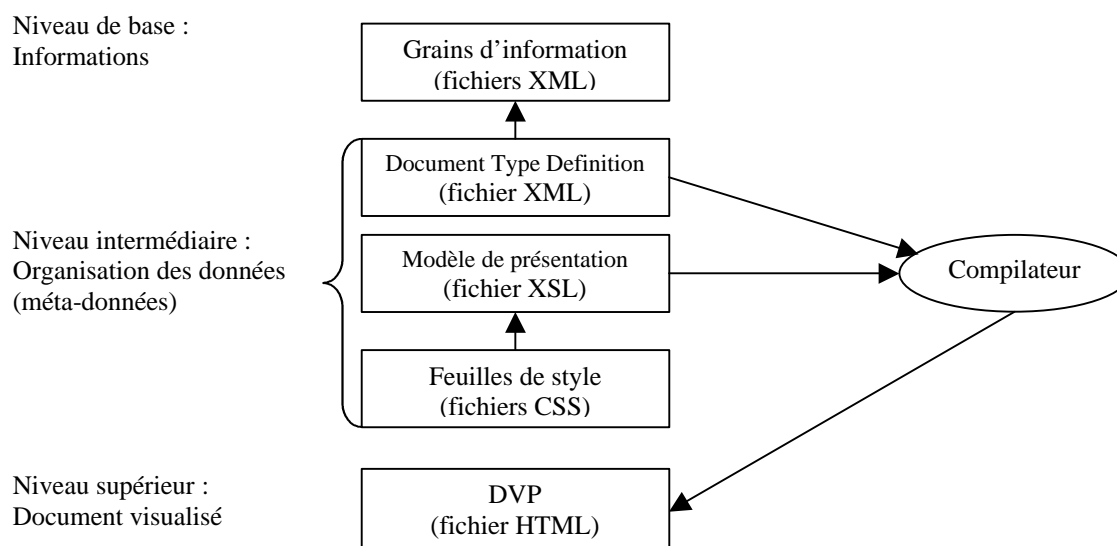
Cette approche, si elle est souvent associée aux technologies XML et XSL décrites au paragraphe précédent, n'est cependant pas nouvelle dans son principe. Comme nous l'avons signalé au chapitre 2, elle hérite de travaux plus anciens réalisés grâce à d'autres moyens technologiques. Par exemple, l'idée de modèle de document est développée en 1989 dans les

⁷² <http://www.w3c.org>. Signalons de plus l'ouvrage en français sur XML grâce auquel nous avons pu effectuer la majorité des développements : [Michard 99].

« documents conceptuels » du système MacWeb et utilise des scripts en langage WebTalk pour spécifier notamment les règles de présentation des modèles de documents ([Nanard & Nanard 89], via [Balpe & al. 96], p.62). Plus récemment, les travaux sur la réutilisation des structures de données, notamment grâce aux « design patterns » de [Garzotto & al. 95] sont repris dans [Nanard & Nanard 98] en utilisant le langage SGML (ancêtre de XML) et la validation des modèles de documents par des DTD.

Cette approche propose un cadre technique pour la conception du LMCE : la génération de DVP nécessite la mise en place d'éléments de conception très proches de ceux créés ou utilisés pour la conception du LMCE, tels qu'une base d'éléments d'information ou des modèles de structure et de présentation. Certains travaux vont même jusqu'à inclure un modèle du domaine d'activité de l'utilisateur pour identifier ses besoins en informations (cf. [Charron & Fluhr 99], [Garlatti & Iksal 99]). De plus, dans la perspective d'une génération automatique des LMCE, l'utilisation de cette approche pour la mise au point de la maquette facilite l'automatisation de la technique de production des LMCE en permettant de séparer les éléments spécifiques aux projets des composants réutilisables.

Le diagramme suivant, issu de [Moulin 99], indique la technique la plus courante qui sous-tend la mise au point de DVP.



Dans ce schéma, il apparaît que la technologie XML/XSL est utilisée pour la composition de documents virtuels personnalisables. Les contenus des fichiers XML et XSL sont transmis à un « compilateur » (encore appelé *processeur* ou *parser XML*) qui n'a pour charge que de « traduire » le document produit en HTML dynamique visualisable. L'appel de ce processeur se fait par l'utilisation de scripts dans les pages HTML. Depuis sa version 5, Microsoft Internet Explorer est livré avec un parser XML spécifique : Msxml. Les éléments de code XML ainsi que les scripts donnés au chapitre suivant ont été testés avec cet outil.

Le prototype de LMCE présenté plus loin est conçu pour fonctionner avec ce processeur. Ainsi, à la différence du prototype de Jean-Marc Charlot, les pages HTML visualisées sont générées dynamiquement à partir des données correspondantes et n'ont pas besoin d'être « pré-compilées ». Par ailleurs, contrairement au processeur utilisé par Jean-Marc Charlot pour son prototype, le processeur Microsoft reconnaît différents types d'encodage de fichiers XML. Nous utilisons ici un encodage conforme à la norme ISO 8859, qui permet l'utilisation d'un jeu de caractères étendu (caractères accentués, cédilles...)

3.3.3. La technologie ActiveXâ

Cependant, pour des raisons de sécurité, ces langages dédiés à la publication sur Internet sont nécessairement limités. Même avec leurs extensions classiques (JAVA, langages de scripts), certaines fonctionnalités sont interdites, comme notamment la lecture ou l'écriture sur le disque de la machine cliente. Ces fonctionnalités peuvent cependant être requises pour le

fonctionnement du LMCE⁷³. C'est pourquoi nous avons recours à la technologie ActiveX®, qui permet de télécharger, d'installer et d'exécuter toutes sortes d'applications dynamiques à partir d'un navigateur web. Le LMCE étant destiné à être distribué sur un intranet (i.e. depuis un serveur interne à l'organisation), les problèmes de sécurité ne se posent pas, et nous pouvons donc diffuser ce type d'applications.

Nous avons choisi de privilégier la facilité de diffusion au détriment de la portabilité. En effet, ce type de contenu est spécifique aux environnements PC sous Windows® et ne fonctionne sur aucun autre type de plate-forme. De plus, après une expérience malheureuse avec Netscape Navigator®⁷⁴, nous avons également choisi de nous limiter à des développements pour le navigateur Internet Explorer® de Microsoft®.

Ces choix technologiques sont naturellement inenvisageables pour le développement d'un système finalisé. Ils sont toutefois acceptables dans le cadre du développement d'un prototype. De plus, l'unité où nous avons mené le projet-pilote et à laquelle nous destinions ce prototype ne disposait que de l'environnement et du navigateur Microsoft®.

3.3.4. Le format pdf et la suite Acrobat®

Comme le format *Postscript*, le format *Portable Document Format* (pdf) a été développé par la société Adobe, initialement pour préserver les formats de sortie des documents. Il permet notamment l'encapsulation des polices de caractères utilisées. Ce format permet de plus la superposition de plusieurs couches de documents (en général une image et du texte), pour conserver la mise en page originelle. Il intègre en outre une compression des données du fichier originel. Ce format est notamment connu pour permettre la diffusabilité des documents en les « figeant » dans une mise en forme définitive⁷⁵. C'est cette particularité qui a guidé le choix de l'unité où nous avons mené le projet-pilote. En effet, dans un domaine sensible comme le nucléaire, les documents émis doivent être les moins modifiables possible.

⁷³ Notamment pour l'installation d'un plug-in dont nous parlons au paragraphe suivant.

⁷⁴ La version 4.7 de Netscape Navigator® dont nous disposions à l'époque n'intégrait pas la technologie ActiveX®. Il fallait, pour pouvoir exécuter des composants ActiveX®, télécharger un plug-in. Ce plug-in nécessitait une liaison Internet pour fonctionner, ce dont nous nous sommes aperçu au cours de la démonstration du prototype à la conférence ICNC'99 ([Médini & al. 99]). Nous avons dès lors décidé de cesser tout développement pour ce navigateur.

⁷⁵ Bien entendu, ces documents peuvent toujours être modifiés à l'aide de la suite complète Acrobat® et de quelques plug-ins spécifiques. Ils sont cependant figés pour les postes sur lesquels n'est installé que le logiciel de visualisation Acrobat Reader® ou même pour ceux qui ne disposent que de la version standard de la suite payante d'Acrobat®.

Adobe vend la suite logicielle Acrobat® (anciennement Acrobat Exchange) qui permet de créer des documents au format pdf, mais distribue gratuitement le logiciel de visualisation Acrobat Reader®. Ces deux applications ont des fonctionnalités de base qui peuvent être étendues grâce à des *plug-ins*. Il s'agit de programmes dédiés à ces applications, qui sont chargés au démarrage, et permettent notamment d'ajouter des fonctionnalités au traitement du document courant ou d'interfacer Acrobat® avec d'autres applications.

Notons que l'utilisateur peut développer ses propres plug-ins pour la version payante d'Acrobat®⁷⁶, mais ne peut installer que des plug-ins possédant un numéro de licence d'Adobe sur Acrobat Reader®. Il est cependant possible de se faire attribuer un tel numéro afin de pouvoir diffuser un plug-in⁷⁷.

Parmi les plug-ins disponibles pour Acrobat® et Acrobat Reader®, le plug-in Acrobat Search® (distribué en standard depuis la version 3.01 et jusqu'aux versions 4 ; il est en option sur la récente version 5), permet la recherche plein texte et par champs dans un corpus de documents pdf. La technologie sous-jacente est le moteur de recherche Topic® de Verity dont le mode d'interrogation, au départ booléen, peut être étendu au langage VDK de Verity. L'outil d'indexation correspondant est disponible dans le plug-in Acrobat Catalog®, livré en standard avec la suite complète Acrobat®.

Le prototype présenté au chapitre suivant utilise les trois types de technologies décrites ci-dessus. Ce prototype s'appuie sur le projet-pilote présenté dans la partie suivante.

3.4. Le projet-pilote

Cette thèse s'appuie sur un projet opérationnel qui a fourni la modélisation nécessaire à la mise au point d'un livre de connaissances électronique.

3.4.1. *L'unité où a été mené le projet*

Le prototype de LCE de Jean-Marc Charlot et le prototype de LMCE proposé dans cette thèse se fondent sur un projet-pilote mené au Service d'Études de Criticité (SEC) de l'Institut de Protection et de Sûreté Nucléaire (IPSN). Le SEC est la référence française en matière de sûreté-criticité nucléaire (voir paragraphe suivant). C'est de plus une unité qui fonctionne au

⁷⁶ Toute la documentation – ainsi que le Standard Development Kit (SDK), qui fournit les outils nécessaires aux développeurs – sont d'ailleurs disponibles gratuitement sur les sites d'Adobe.

sein du CEA depuis près de cinquante ans, et qui depuis lors s'est créé un système d'information imposant et d'une grande diversité. Ces deux raisons en font un exemple d'application idéal pour notre étude. Le SEC, de son côté, éprouve un réel besoin de structurer son patrimoine d'informations et de connaissances, d'une part pour assurer sa mission de diffusion des connaissances en criticité auprès des acteurs de la filière nucléaire française, et d'autre part pour continuer à faire progresser rapidement la recherche dans ce domaine.

3.4.2. *Le domaine de connaissances*

La matière fissile (uranium, plutonium...) émet et absorbe en permanence des neutrons. Lors des réactions en chaîne dans les réacteurs nucléaires, le rapport des nombres de neutrons émis et absorbés, appelé K_{eff} (ou K effectif) est égal à 1. Cette valeur dépasse l'unité dans le cas des réactions non contrôlées. On dit alors que la matière fissile atteint son seuil de criticité. Lors du transport, de l'entreposage et de la transformation de matières fissiles, il faut pouvoir garantir la sous-criticité de ces matières ($K_{\text{eff}} < 1$). Savoir assurer la sous-criticité de la matière fissile dans toutes ces conditions est le domaine de la maîtrise du risque de criticité hors réacteur nucléaire, aussi appelé sûreté-criticité nucléaire.

Plusieurs moyens permettent de contrôler cette sous-criticité. Un mode de contrôle est l'un de ces moyens de limitation du K_{eff} . Le volume ou la masse limitant la quantité de matière fissile regroupée ou la concentration d'une solution contenant de la matière fissile sont des modes de contrôle couramment employés en sûreté-criticité.

3.4.3. *La modélisation des connaissances*

Le projet de Gestion des Connaissances que nous avons mené au SEC a comporté au total 100 heures d'entretiens, auprès de 7 experts et a permis de produire une centaine de pages de diagrammes de connaissances, du type de celui présenté ci-dessous.

Il avait déjà été mené, en 1995, un début de projet de gestion des connaissances sur la maîtrise du risque de criticité nucléaire, qui avait abouti à un modèle du domaine et un début de modèle d'activité. Du fait des emplois du temps chargés des protagonistes, ces modèles n'ont pas pu être validés, et le projet n'a pu être achevé. Nous avons bien entendu conservé les modèles existants en reprenant ce projet et les avons soumis aux experts pour validation. Le modèle du domaine, qui qualifie l'ensemble des processus physiques en jeu dans le

⁷⁷ De ce fait, la diffusion et la vente de plug-ins pour Acrobat sont courantes. Elles sont assurées soit par des sociétés appartenant au « club des développeurs Acrobat », soit sur des sites dédiés, comme « planet PDF » ou

phénomène de la criticité nucléaire n'a que très peu été modifié. Ce modèle est, à notre avis, un peu général par rapport au degré de finesse de la modélisation effectuée ensuite dans chacun des sous-domaine. Cependant, il est nécessaire pour présenter toutes ces activités par rapport aux différents aspects du phénomène qu'elles étudient. Le modèle d'activité, qui ne concernait qu'un des sous-domaines d'activité du SEC a servi de base à la modélisation dans ce sous-domaine, et a été considérablement amélioré et étoffé.

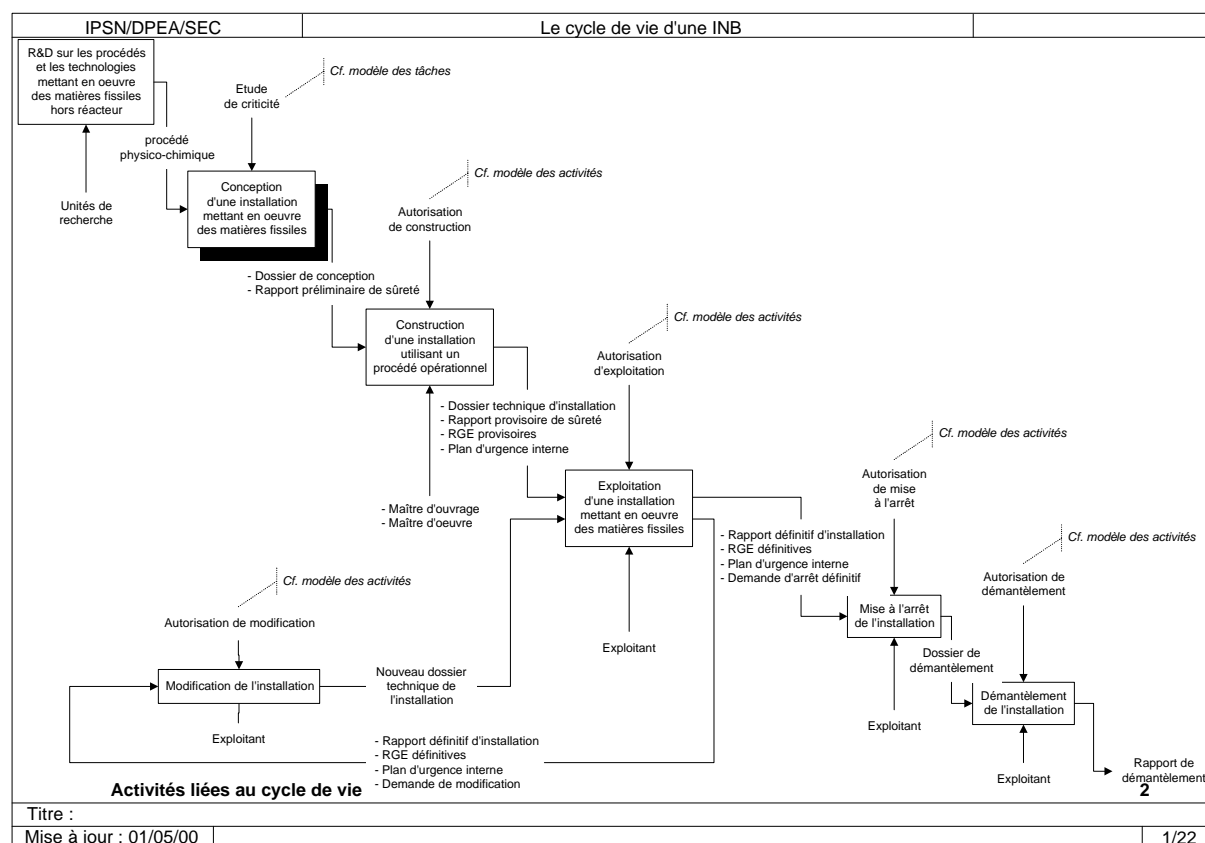


Figure 21. Exemple de diagramme de connaissances en sûreté-criticité nucléaire dans le sous domaine des études de criticité pour les INB.

Six sous-domaines d'activité ont été identifiés au SEC. Ces sous-domaines sont respectivement regroupés deux à deux en trois grandes classes : la recherche et le développement, l'ingénierie et le rôle de référence française en criticité. Seuls les deux premières ont donné lieu à une modélisation des connaissances satisfaisantes. La liste de ces sous-domaines et des modèles de connaissances qui ont été réalisés dans chacun d'eux est donnée dans le tableau ci-dessous.

Domaine ou sous-domaines	Nombre de diagrammes réalisés			
	Activités	Tâches	Concepts	Objets
Développement, qualification et maintenance de codes de calculs en criticité	6	12	2	2
Participation aux programmes d'expériences en criticité,	7	9	2	0
Études de criticité pour les exploitants d' <i>Installations Nucléaires de Base</i> (INB)	12	9	6	2
	(+ 1 document textuel)			
Expertises en criticité dans le domaine des transports de matières fissiles pour l'autorité nucléaire française	4	9	1	0
	(+ 1 document textuel)			
Diffusion des informations et des connaissances en criticité à la communauté nucléaire française	2	0	0	0
Établissement de standards et de normes en criticité	0	0	0	0
Sûreté-criticité nucléaire	7	Modèle du domaine : 15		

Tableau 3. Nombre de diagrammes réalisés par sous-domaines.

3.4.4. Inventaire du système d'information

Ses activités conduisent le SEC à gérer une quantité toujours croissante de documents, recueils de données, outils de traitement... qui en constituent le système d'information. L'ampleur et la disparité de celui-ci en rendent l'utilisation complexe.

En parallèle avec notre projet, le SEC a mené une action de numérisation de ses archives. Plus de 50 000 pages ont été numérisées et sont désormais disponibles au format *pdf* d'Adobe.

Nous avons effectué un inventaire des éléments de ce SI et en avons élaboré un modèle spécifique au domaine. Les modèles obtenus sont présentés aux paragraphes 4.2.1.3.1 et 4.3.1.3 du chapitre suivant..

3.5. Bilan des perspectives méthodologiques

Dans ce chapitre, nous avons présenté les différents outils et méthodes sur lesquels s'appuie ce travail. Les deux disciplines détaillées sont la gestion des connaissances et l'ergonomie. Le principe et les différentes techniques employées pour réaliser l'outil présenté au chapitre suivant sont décrites. Le domaine d'application du projet-pilote dans lequel est réalisé le prototype est présenté.

La méthode de gestion des connaissances MKSM est une méthode de modélisation qui permet d'obtenir une vision structurée du contexte d'activité d'une organisation. Cette thèse se place dans une perspective d'opérationnalisation de cette méthode pour l'accès à l'information. Pour cela, nous utilisons cette méthode telle qu'elle est définie par ses concepteurs pour la modélisation du contexte d'activité de cet accès à l'information. Nous avons défini une méthode de modélisation des utilisateurs spécifique pour notre approche. Cette méthode est issue d'une synthèse de différents travaux en ergonomie et en recherche d'information. Elle est évoquée plus en détails au chapitre 4. La méthode de modélisation du système d'information est également décrite dans le chapitre suivant, car elle résulte des particularités des deux modèles ci-dessus.

Concernant la méthode de conception de l'outil, l'ergonomie des logiciels propose différentes techniques. Parmi celles-ci, nous avons sélectionné celles qui nous paraissent pertinentes pour la conception d'un système hypermédia d'accès à l'information par recontextualisation. Nous avons également choisi d'utiliser la méthode MKSM comme référentiel pour les différents éléments de cette phase de conception. La méthode de conception logicielle employée au chapitre suivant est issue de ces choix. Elle est donc spécifique à notre approche et fait l'objet d'une discussion a posteriori au chapitre 5. L'un des objectifs méthodologiques de cette thèse est d'identifier les modalités et la pertinence d'une méthode de conception d'applications justifiée et fondée sur une modélisation multi-points-de-vue du domaine cible de ces applications.

Bibliographie

- [Addis 96] : Addis, T. R., Addis, J. J. T., "A Functional Schema Interpreter as an aid to Knowledge System Design", in proceedings of the Thinking with Diagrams colloquium, IEE Computing and Control Division, Professional Group C4 (Artificial Intelligence), Digest No: 96/010, pp. 13/1-13/10, London, 18 janvier 1996.
- [Allen 97] : Allen R. B., "Mental Models and User Models", Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu Eds., Elsevier Science, 1997, pp.49-63.
- [Balpe & al. 96] : Balpe J-P., Lelu A., Papy F., Saleh I., "Techniques avancées pour l'hypertexte", Hermès, Paris 1996.
- [Barthes 64] : Barthes R., "Éléments de sémiologie", 1964.
- [Bastien & Scapin 93] : Bastien, J. M. C., Scapin D-L., "Ergonomic Criteria for the Evaluation of Human-Computer Interfaces", Rapport technique INRIA n° 156, Juin 1993.
- [Bastien & al. 98] : Bastien J. M. C., Leulier C., Scapin D-L., "L'ergonomie des sites web", in J-C. Le Moal & B. Hidoine Eds., Créer et maintenir un service web, Paris, ADBS, pp. 111-173, 1998.
- [Booch 92] : Booch G., "Conception orientée objets et applications", Addison-Wesley, 1992.
- [Buckingham Shum 96] : Buckingham Shum S., "Design Argumentation as Design Rationale", in The Encyclopaedia of Computer Science and Technology, Kent A. and William J. G. (Eds), Marcel Dekker Inc., vol. 35, Suppl. 20, pp. 95-128, New York, disponible en ligne : <ftp://kmi-ftp.open.ac.uk/pub/simonb/ecst96.pdf>, 1996.
- [Buckland 91] : Buckland M., "Information as thing", journal of the American Society of Information Science, vol. 42, N° 5, pp. 351-360, disponible en ligne : <http://www.sims.berkeley.edu/~buckland/thing.html>, 1991.
- [Bush 45] : Bush V., "As we may think", Atlantic Monthly, 176 (1), pp. 101-108, 1945.
- [Chaillot & Ermine 97] : Chaillot M., Ermine J-L., "Le livre de connaissances électronique", Document numérique, vol. 1, n°1, 75-98, 1997.
- [Chang & Rice 93] : Chang S-J., Rice R-E., "Browsing: a multidimensional framework", Annual review of information science and technology, Vol.28, Learned Inf., Medford, NJ, USA; xv, 1993, pp. 231-76.
- [Charlot & al. 00] : Charlot J-M., Chaillot M., Médini L., "Information Engineering based on knowledge modelling: a case study in nuclear criticality safety", International Seminar on Integrated Information Systems, IAEA-SR-212/46, Vienne (Austria), april, 10-14, 2000.
- [Charreton & Ermine 96] : Charreton B., Ermine J-L., "From knowledge specification to executable specification", in proceedings of KEML'96, Knowledge Engineering and Modelling Languages, Paris, 1996.
- [Charron & Fluhr 99] : Charron J., Fluhr C., "BeFor (Beyond Forms). Un modèle de représentation du Web invisible", Actes de la conférence H²PTM'99, Hermès, pp. 27-41, Paris, 1999.
- [Claverie 97] : Claverie B., Le Blanc B., Corsini M., "Étude des processus cognitifs privilégiés dans l'exploitation des bases de connaissances pour une navigation dynamique liée à l'utilisateur", sujet de thèse proposé à l'université de Bordeaux, 1997.
- [de Saussure 16] : de Saussure F., "Cours de linguistique générale", 1916.

- [de Rosnay 75] : de Rosnay J., *"Le macroscopie : vers une vision globale"*, Paris: Seuil, 1975.
- [DISTNB 98] : Direction de l'information scientifique, des technologies nouvelles et des bibliothèques (DISTNB), Département de l'information spécialisée, Définition de l'IST disponible en ligne : <http://www.mesr.fr/icst/ist/index.htm>, 1998.
- [Ermine & al. 96] : Ermine J-L., M. Chaillot, P. Bignon, B. Charreton, D. Malavieille: *"MKSM, a method for knowledge systems management"*. in Knowledge Management, Organization, Competence and Methodology, Advances in Knowledge Management, Volume 1, Jos F. Schreinmakers Ed., pp. 288-302, Ergon Verlag, 1996.
- [Ermine 97] : Ermine J-L., *"MKSM, méthode pour la gestion des connaissances"*, polycopié de cours n° 9, DEA SETI, CEA Saclay, 1997.
- [Farenc & Palanque 99] : Farenc C., Palanque P., *"Exploitation des notations de Design Rationale pour une conception justifiée des applications interactives"*, in Interaction Homme-Machine, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Cépaduès-Éditions, Toulouse, France, vol. 1, pp. 33-40, 1999.
- [Felber 87] : Felber H., *"Manuel de terminologie"*, Paris, UNESCO, 1987.
- [Finke 89] : Finke R. A., *"Principles of mental imagery"*, MIT Press, Cambridge, Mass., 1989.
- [Fisher & al. 96] : Fisher G., Lemke A. C., Mc Call R., Morch A. I., *"Making Argumentation Serve Design"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp.393-419, réédité dans Moran T. P. & Carroll J. M. (Eds.), Design Rationale: Concepts, Techniques and Use, Lawrence Erlbaum Associates, pp. 267-293, 1996.
- [Foucault 99] : Foucault B., *"Analyse des stratégies de navigation sur le Web : expériences sur les modes d'accès à l'information"*, in Interaction Homme-Machine, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Cépaduès-Éditions, Toulouse, France, vol. 2, pp. 45-48, 1999.
- [Fucella & al. 98] : Fucella j., Pizzolato J., Franks J., *"Web site user centered design: technique for gathering requirements and tasks"*, Internetworking, 1 (1), disponible en ligne : http://www.sandia.gov/itg/newsletter/june98/web_design.html, 1998.
- [Fucella & Pizzolato 98] : Fucella j., Pizzolato J., *"Creating Web site design based on user expectations and feedback"*, Internetworking, 1 (1), disponible en ligne : http://www.sandia.gov/itg/newsletter/june98/user_requirements.html, 1998.
- [Garlatti & Iksal 99] : Garlatti S., Iksal S., *"Documents virtuels personnalisables pour les systèmes d'information"*, in Interaction Homme-Machine, Actes de l'Atelier sur les Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, pp. 22-26, Novembre 1999.
- [Garzotto & al. 95] : Garzotto F., Paolini P., Mainetti L., *"Hypermedia Design, Analysis and Evaluation Issues"*, in Communications of the ACM, n° 38, vol. 8, pp. 74-86, Août 1995.
- [Glasgow & al. 95] : Glasgow J., Narayanan, H., Chandrasekaran B., *"Diagrammatic Reasoning, Cognitive and Computational Perspectives"*, The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, London, England, 1995.
- [IGL 82] : IGL France, *"Introduction à SADT"*, manuel, 1982.
- [Johnson-Laird 83] : Johnson-Laird P. N., *"Mental models"*, Harvard University Press, Cambridge, Mass., 1983.
- [Kass & Finin 88] : Kass T., Finin T., *"The Need for User Models in Generating Expert*

- System Explanations*", in International Journal of Expert Systems Research and Applications, vol. 1., pp. 345-375, 1988.
- [Kosslyn 80] : Kosslyn S. M., *"Image and Mind"*, Harvard University Press, Cambridge, Mass., 1980.
- [Kristen 98] : Kristen G., transparents issus de la présentation de la méthode KISS, organisée au CEA de Saclay en avril 1998.
- [Kuntz & Rittel 70] : Kuntz W., Rittel H. W. J., *"Issues as elements of information systems"*, Working paper No. 131, Berkeley, UCB, Center for Planning and Development Research, 1970.
- [Lainé-Cruzel 94] : Lainé-Cruzel S., *"Vers de nouveaux systèmes d'information prenant en compte le profil des utilisateurs"*, Documentaliste – Sciences de l'information, vol. 31, n° 3, pp. 143-147, 1994.
- [Lee & Lai 91] : Lee J., Lai K., *"What's in Design Rationale?"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp.251-280, réédité dans Moran T. P. & Carroll J. M. (Eds.), *Design Rationale: Concepts, Techniques and Use*, Lawrence Erlbaum Associates, pp. 125-154, 1996.
- [Le Moigne 77] : Le Moigne J.L., *"La théorie du système général, Théorie de la modélisation"*, Paris, PUF, 1977, 3^{ème} édition mise à jour, 1990.
- [Leulier & al. 98] : Leulier C., Bastien J. M. C., Scapin D-L., *"Compilation of ergonomic guidelines for the design and evaluation of websites"*, in Commerce and Interaction Report, INRIA, Rocquencourt, France, 1998.
- [Lippold & al. 94] : Lippold B., Pomian J., Henry, J. Y., Elsensohn O., *"AVIS : une méthode d'analyse de la cohérence des documents"*, in proceedings of the European Safety and Reliability Conference, pp. 888-899, La Baule, 1994.
- [Mackay 88] : Mackay, W. E., *"More than Just a Communication System: Diversity in the Use of Electronic Mail"*, Proceedings of CSCW'88: Conference on Computer-Supported Cooperative Work, Portland, Oregon: ACM, 1988.
- [Malavieille & al. 95] : Malavieille D., Ermine J-L., Ct Ribot, Lions C., Lt Condamin, *"Étude et réalisation d'un système à base de connaissances d'aide à la décision en cas d'intervention sur des accidents majeurs"*, in proceedings of IA'95, Montpellier, 1995.
- [Marchionini 95] : Marchionini G., *"Information seeking in electronic environments"*, New York, Cambridge University Press, 1995.
- [Mc Call 79] : Mc Call R. *"On the structure and use of issues systems in design"*, unpublished doctoral dissertation, Berkeley, UCB, 1979.
- [Mc Knight & al. 91] : Mc Knight, C., Dillon A., Richardson J., *"Hypertext in context"*, Cambridge University Press, Cambridge, 1991.
- [Mc Lean & al. 91] : Mc Lean A., Young R. M. Bellotti V., Moran T., *"Questions, Options, and Criteria: Elements of Design Space Analysis"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp. 201-250, réédité dans Moran T. P. & Carroll J. M. (Eds.), *Design Rationale: Concepts, Techniques and Use*, Lawrence Erlbaum Associates, pp. 75-124, 1996.
- [Médini 97] : Médini L., *"Conception et Réalisation d'un Livre de Connaissances Électronique sous forme de Document Hypermédia"*, Rapport de projet de DEA, Orsay et INSTN, Sept. 1997.

- [Médini & al. 99] : Médini L., J. Anno, M. Chaillot, V. Rouyer, D. L. Scapin, *"Towards a knowledge oriented system for information management in nuclear safety"*, Proceedings of the Sixth International Conference on Nuclear Criticality-Safety, Versailles, Sept. 1999, 1451-1461.
- [Michard 99] : Michard A., *"XML, Langage et applications"*, Eyrolles, Paris, 1999.
- [Millerat & al. 96] : Millerat P., Ermine J-L., Chaillot M., *"Knowledge management for modelling nuclear power plants control in incidental and accidental states"*, Proceedings of the CESA'96 IMACS Multiconference, Symposium on Modelling, Analysis and Simulation, Vol. 2, pp.982-987, Lille, 1996.
- [Moran & Carroll 96] : Moran T. P., Carroll J. M. (Eds.), *"Design Rationale: Concepts, Techniques and Use"*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, ISBN : 0-8058-1566-X, 1996.
- [Moulin 99] : Moulin C., Pazzaglia J-C., *"Documents pédagogiques adaptatifs dans un environnement d'apprentissage distribué"*, in Interaction Homme-Machine, Actes de l'Atelier Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, pp. 27-31, Novembre 1999.
- [Nanard & Nanard 89] Nanard M., Nanard J., *"MacWeb, un outil pour élaborer des documents"*, Ed. Bigre, Rennes, pp. 18-29, 1989.
- [Nelson 67] : Nelson T. H., *"Getting it out of our system"*, in Information Retrieval: a Critical Review, G. Schechter (Ed.), Thompson Books, Washington D. C., 1967.
- [Nonaka 95] : Nonaka I., Takeuchi H., *"The knowledge-creating company"*, Oxford University Press, 1995.
- [Normann 88] : Normann D. A., *"The Psychology of Everyday Things"*, Basic Books, Inc, New York, 1988.
- [Panijel 99] : Panijel C., Unité régionale de promotion et de formation pour l'information scientifique et technique de Paris / École des Chartes, Information Scientifique et Technique, disponible en ligne : <http://www.ccr.jussieu.fr/urfist/def-ist.htm>, 1999.
- [Picard 96] : Picard S., *"Conception et réalisation d'un moteur abstrait hypermédia pour la méthode MKSM"*, Orsay et INSTN, septembre 1996.
- [Pomian 96] : Pomian J., *"Mémoires d'entreprises, techniques et outils de la gestion du savoir"*, Paris, Sapiencia, 1996.
- [Scapin 86] : Scapin D-L., *"Guide ergonomique de conception des interfaces homme-machine"*, Rapport Technique INRIA n° 77, Rocquencourt, Octobre 1986.
- [Scapin 97] : Scapin D-L., *"Ergonomie des logiciels Interactifs"*, transparents issus du cours d'ergonomie du DEA SETI, CEA Saclay, 1997.
- [Scapin & Pierret-Golbreich 90] : Scapin D-L., Pierret-Golbreich C., *"Towards a method for task description: MAD"*, in L. Berlinguer and D. Berthelette (Eds), Work in display units 89, Elsevier Science Publishers, North Holland, 1990.
- [Schreiber & Wielinga 93] : Schreiber G., Wielinga B., *"Model construction"*, in Schreiber G., Wielinga B., Breuker J. (Eds), *"Kads, a principled approach to knowledge-based systems development"*, Academic Press, 1993.
- [Sector 98] : notes personnelles prises lors de la visite de l'entreprise Sector, dans le cadre de la formation « Les Doctoriales », Juin 1998.

- [Shannon 48] : Shannon C. E. *"A mathematical Theory of Communication"*, The Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, juillet et octobre 1948.
- [Tardieu & al. 83] : Tardieu H., Rochefeld A., Coletti C., *"La méthode Merise"*, Éditions d'organisation, 1983.
- [Van Campenhoudt 96] : Van Campenhoudt M., *"Abrégé de terminologie multilingue"*, centre de recherche Termisti, disponible en ligne : <http://www.refer.fr/termisti/theoweb3.htm>, 1996.
- [Van Craeynest & al. 00] : Van Craeynest J-M., Charlot J-M., Malavieille D., Bigeon P., Chaillot M., *"Traçabilité des connaissances avec MKSM"*, Actes du congrès scientifique NîmesTIC'2000, Nîmes, pp. 12-22, Septembre 2000.
- [Vogel 88] : Vogel. C., *"Génie cognitif"*, Collection Sciences Cognitives, Masson, Paris, 1988.
- [Von Bertalanffy 1968] : Von Bertalanffy L., *"General Systems Theory"*, Foundation, Development, Applications, New York: G. Braziller (1968), Trad. française: *"Théorie générale des systèmes"*, Paris: Dunod, 1973.
- [Von Hippel 88] : Von Hippel, E., *"The Sources of Innovation"*, New York, NY: Oxford University Press, 1988.
- [Vora 98] : Vora P., *"Human factors methodology for designing web sites"*, in Human factors and web development, Forsythe C., Grose E. & Ratner J. (Eds.), Mahwah, New Jersey, Lawrence Erlbaum Associates, pp. 153-172, 1998.
- [Waern & Hägglund 97] : Wærn, Y., Hägglund S., *"User Aspects Of Knowledge-Based Systems"*, Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu (eds.), Elsevier Science, 1997, pp. 1159-1176.
- [Wahlster & Kobsa 89] : Wahlster W., Kobsa A., *"User Models in Dialog Systems"*, A. Kobsa and W. Wahlster eds., Berlin, Springer-Verlag, pp. 4-34, 1989.
- [Wielinga & al. 93] : Wielinga B., Schreiber G., Breuker J., *"Modelling expertise"*, in Schreiber G., Wielinga B., Breuker J. (Eds), *"Kads, a principled approach to knowledge-based systems development"*, Academic Press, 1993.
- [Woodhead 90] : Woodhead N., *"Hypertext & Hypermedia. Theory and Applications"*, Sigma Press, GB, 1991.
- [Wüster 76] : Wüster E., *"La théorie générale de la terminologie -un domaine interdisciplinaire impliquant la linguistique, la logique, l'ontologie, l'informatique et les sciences des objets "*, Actes du colloque international de terminologie (Québec, Manoir du lac Delage, 5-8 octobre 1975) , Essai de définition de la terminologie, H. DUPUIS éd., Québec, Régie de la langue française, p. 49-57, 1976.

Chapitre 4. Contribution de la thèse : l'identification et les éléments de solution proposés pour la problématique de l'accès à l'information

4.1. Problématique

La problématique générale de cette thèse concerne l'élaboration d'une stratégie spécifique d'accès à l'information à partir de la modélisation d'un domaine d'activité et la définition de la méthodologie de conception de l'architecture logicielle générique correspondant à cette stratégie. Cette problématique pluridisciplinaire comporte deux types de difficultés. D'une part, il s'agit de considérer chacune de ces disciplines avec un niveau de détails suffisant mais cependant limité par rapport à cette problématique. D'autre part, l'objectif est de les articuler selon un raisonnement qui met au jour l'intégration de chacune d'elles dans un processus global de conception.

Nous abordons ici cette problématique en plusieurs parties qui nous permettent de détailler les différents aspects sur lesquels porte ce travail de thèse. Cela nous permet ensuite de présenter ce travail en articulant chaque partie de façon cohérente par rapport à la problématique d'ensemble.

Les deux premières parties concernent les aspects méthodologiques de ce travail. Il s'agit d'une part de formaliser et d'articuler les éléments de modélisation sur lesquels se fonde la structure conceptuelle du livre de connaissances électronique et d'autre part de mettre au point la stratégie d'accès à l'information qui particularise cet outil. Dans un deuxième temps, nous présentons la problématique opérationnelle de conception d'un système d'accès à l'information répondant à toutes ces contraintes méthodologiques. Enfin, nous nous intéressons à l'aspect pluridisciplinaire de ce type d'outil, et plus particulièrement au statut des modèles diagrammatiques par rapport à une démarche d'accès à l'information. Cette dernière partie est abordée au chapitre 5 et permet d'introduire la discussion sur le bien-fondé de l'approche d'accès à l'information avec le livre de connaissances électronique.

4.1.1. *Méthodologie de modélisation*

Nous avons vu aux chapitres précédents l'intérêt de disposer d'une modélisation multi-points-de-vue pour recontextualiser l'information dans son domaine d'utilisation : un outil d'accès à l'information adoptant une telle démarche permet à l'utilisateur d'identifier les documents pertinents par rapport à chaque aspect de son activité. Toutefois, la richesse de la modélisation multi-points-de-vue a pour but de présenter une vision globale de l'activité d'une organisation. Elle diffère des spécificités des problèmes d'information (tâche nécessitant une recherche d'information, niveau de compréhension de cette information, connaissance de l'outil) des individus qui l'utilisent dans le cadre d'une démarche d'accès à l'information. En outre, une démarche d'accès à l'information à partir d'une modélisation « orientée-métier » de l'activité de l'organisation requiert également une vision orientée-métier de son système d'information.

Il nous semble donc que ces spécificités doivent être prises en compte par ailleurs. C'est pourquoi nous formulons l'hypothèse qu'une démarche d'accès à l'information fondée sur la recontextualisation de cette information dans son contexte d'utilisation ne peut s'appuyer uniquement sur des modèles de connaissances. Nous avons donc fondé cette démarche sur une approche multi-modèles, incluant, en plus du modèle de l'activité de l'organisation (encore appelé *modèle du domaine*), des modèles des utilisateurs et du système d'information.

Pour chacun des deux derniers modèles, deux questions sont posées : quelle forme lui donner et comment l'obtenir ? L'explicitation de la première question pour chacun des deux modèles est spécifique à ce modèle. Il s'agit de préciser leurs particularités et les éléments qu'ils doivent nécessairement contenir pour permettre l'accès à l'information. Un élément de réponse à la seconde question est commun à ces deux modèles et est développé plus loin.

De même qu'en gestion des connaissances, les méthodes de modélisation de l'information (i.e. d'indexation documentaire) ou d'utilisateurs ne manquent pas. Cependant, deux impératifs nous conduisent à définir des modèles spécifiques.

Dans un souci de rentabilité, il est souhaitable que la phase d'interview des experts dure le moins longtemps possible. En effet, c'est la phase la plus coûteuse de la modélisation, car elle monopolise deux individus (dont l'expert, dont le temps est par définition précieux pour l'organisation), alors que la mise en forme des modèles à partir des traces de l'interview ne mobilise que le chargé de projet. Or, lors de la phase de modélisation des connaissances, le *chargé de projet en Gestion des Connaissances* est directement au contact des futurs

utilisateurs du système et de l'information disponible dans l'organisation. Il est donc logique d'envisager d'utiliser cette étape de modélisation pour recueillir les informations nécessaires pour mettre au point les deux autres modèles.

De plus, la modélisation des connaissances obtenue est riche, et comporte, entre autres, une modélisation des tâches qui s'apparente à celles utilisées en ergonomie. C'est pourquoi il nous a semblé intéressant d'approfondir les aspects de la modélisation des connaissances qui peuvent être réutilisés pour la mise au point des autres modèles, et notamment celui des utilisateurs.

Les modèles du SI et des utilisateurs sont donc très dépendants de celui de l'activité de l'organisation. À ce titre, ils peuvent être élaborés à partir et en fonction de ce dernier. L'objectif est donc d'identifier la problématique de définition de ces modèles en fonction de celui du domaine. Il s'agit également d'évaluer dans quelle mesure les informations fournies par la modélisation des connaissances doivent être complétées par des recueils de données spécifiques.

La problématique de définition et de mise au point de ces modèles est détaillée – pour chacun d'eux – dans les paragraphes suivants. Il s'agit ensuite de positionner le problème de mise en cohérence de ces trois modèles. Ce positionnement nous permet alors d'aboutir à une vision conceptuelle du LMCE, sur laquelle nous nous fondons pour la démarche de conception de cet outil.

4.1.1.1. Vers la définition d'un modèle de l'utilisateur

Quels que soient les approches et les paradigmes retenus, la majorité des auteurs qui ont travaillé sur les systèmes d'accès à l'information s'accordent pour dire que l'utilisateur est un élément déterminant pour la conception de ces systèmes (voir par exemple [Lainé-Cruzel 94] ou [Palmquist 98] pour les stratégies analytiques, [Nielsen 90] ou [Woodhead 90] pour les hypermédias, [Marchionini 95] ou [Agosti & Smeaton 96] pour la combinaison des deux stratégies ou [Waern & Hagglund 97] pour les systèmes à base de connaissances).

Dans notre approche d'accès à l'information orientée-métier, disposer d'un modèle de l'utilisateur est d'autant plus important que cette approche s'adresse à tous les acteurs de l'organisation modélisée, quels que soient leurs situations de travail et les types d'information scientifique et technique recherchés. Dans le cadre de notre démarche de conception ergonomique d'un système d'accès à l'information, l'intérêt d'un modèle des utilisateurs est

également justifié par le désir de présenter à chacun d'eux une interface qui corresponde à ses préférences de visualisation.

Nous avons vu aux chapitres précédents l'intérêt d'un modèle des utilisateurs à trois niveaux : il faut connaître leurs activités pour pouvoir leur présenter les diagrammes de connaissances pertinents pour les problèmes d'information qu'ils peuvent se poser ; il faut également tenir compte de leurs niveaux de connaissance des systèmes d'accès à l'information de façon à pouvoir leur proposer les stratégies d'accès à l'information adaptée à leurs modes de recherche d'information ; il faut enfin connaître leurs préférences de visualisation pour pouvoir leur proposer une interface utilisable.

La problématique de définition des modèles de l'utilisateur est donc d'identifier et de formaliser les caractéristiques se rapportant à chacun des niveaux ci-dessus. Il s'agit de bâtir des *profils d'utilisateurs* qui permettent à la fois le fonctionnement du système d'accès à l'information et la mise au point de composants d'interface adaptés aux différentes situations rencontrées.

Nous avons vu au chapitre 3 la méthode de définition de ces profils. Cette méthode s'inspire de méthodes génériques ou adapte des méthodes spécifiques issues de la littérature. Elle permet d'aboutir à des profils « conceptuels » des utilisateurs. Il convient alors de discuter les aspects techniques de ces profils, de façon à déterminer des profils « génériques ». Ce sont ces profils génériques qui sont instanciés grâce aux données recueillies lors de la phase de modélisation du projet-pilote.

4.1.1.2. Modélisation du système d'information

Lorsqu'un chargé de projet arrive dans une unité pour y mener un projet de gestion des connaissances, il part en général de zéro. En effet, cette discipline étant relativement nouvelle, nous n'avons pas à nous soucier de l'existant. Le livre de connaissances produit, quelle que soit sa forme, contient alors le résultat de tous les travaux menés en gestion des connaissances. Dans la mesure où il n'interagit pas avec un autre artefact du même type, celui-ci n'a besoin que d'être cohérent avec lui-même.

La problématique de la gestion de l'information est tout autre. L'objectif n'est pas ici de formaliser un objet immatériel mais de rendre accessible l'existant pour aider à la production de nouvelles sources d'information par les individus dans l'unité. Un projet de gestion de l'information a donc pour objectif de « faire avec » l'existant. En particulier, pour qu'un projet de gestion des connaissances soit lancé dans une organisation, il faut que celle-ci ait

préalablement acquis une certaine quantité de savoir faire – et donc un SI conséquent – au cours de son activité passée. Cette activité est liée à un processus de production et de stockage de l'information dans lequel nous n'avons pas à intervenir, puisque la modélisation des connaissances réalisée tient compte de ces processus au début du projet. Notre influence se situe en aval, au niveau de l'utilisation de l'information stockée.

C'est pourquoi nous devons mettre au point une méthode d'indexation spécifique des éléments du système d'information⁷⁸ des unités modélisées. Cette méthode doit être suffisamment générique pour être réutilisable et pouvoir s'accommoder de toutes les sources d'informations existantes⁷⁹. Elle doit également être suffisamment adaptable pour correspondre, pour chaque projet, au modèle spécifique du domaine. Par ailleurs, ce modèle doit être compatible avec les deux autres (i.e. celui du domaine et des utilisateurs). Cette compatibilité signifie notamment que certains descripteurs des éléments du SI doivent être des caractéristiques des éléments des autres modèles à partir desquels une recherche d'information doit permettre de trouver un nombre satisfaisant de réponses pertinentes.

Toutes ces contraintes nous laissent penser qu'une modélisation de type indexation conceptuelle (cf.chapitre 2) est adaptée à notre problématique. En effet, elle présente l'avantage de ne pas dépendre de la structure physique des systèmes d'information modélisés (c'est le propre d'une indexation) et peut être définie en fonction d'autres modèles conceptuels (par exemple ceux du domaine et des utilisateurs). Le travail de mise au point du modèle générique des systèmes d'information consiste donc à définir un type d'indexation conceptuelle aussi proche des deux autres modèles que possible. Ceci dans le but de simplifier la fonction de comparaison documentaire lors de la recherche d'information : nous pouvons ainsi réaliser cette fonction de comparaison avec deux niveaux : un premier filtrage est effectué justement en fonction de la mise en correspondance des modèles d'activité et de l'utilisateur avec celui des sources d'information⁸⁰ ; ensuite, le SRI utilisé sélectionne les

⁷⁸ Nous avons vu au Chapitre 2 que ce type de modélisation correspond à la définition d'une indexation documentaire de ces sources d'information. Nous utilisons donc indifféremment les termes modélisation ou indexation à propos des systèmes d'information dans la suite.

⁷⁹ Nous nous limitons ici à l'information accessible électroniquement. L'information papier, qui ne peut être que référencée et donc mise à disposition automatiquement, pourrait cependant être modélisée de la même façon. Cependant, ce support empêche tout processus d'indexation automatique nécessaire à l'intégration de ces documents dans les bases gérées par le(s) système(s) de recherche d'information utilisé(s), et il n'est pas question, pour le chargé de projet, d'indexer manuellement des documents papier spécifiques au domaine.

⁸⁰ Cette étape aboutit à la sélection d'une liste d'index pertinents par rapport à l'individu et à sa situation de travail. La mise en œuvre de ce filtrage relève de la problématique opérationnelle et sera abordée plus loin.

éléments pertinents du système d'information en appariant les termes de la requête et les index documentaires sélectionnés.

Cette indexation doit s'appuyer sur une méthode de modélisation adaptée. Le chargé de projet qui arrive dans une unité peut par exemple procéder à un inventaire systématique du système d'information et des habitudes d'accès à l'information des différents acteurs de l'organisation. Mais la question qui se pose ici est de faire coexister au mieux ces méthodes avec la phase de modélisation des connaissances. Que faut-il faire en plus des interviews de modélisation des connaissances pour connaître les besoins en information d'un individu dans une situation de travail donnée ? Inversement, comment évaluer si un document est plus utile dans une situation plutôt que dans une autre ? Bien entendu, il est difficile de trouver des réponses simples et génériques à ces questions. Mais notre but est ici d'y apporter des éléments de réponse méthodologiques qui favorisent une indexation du SI adaptée aux contraintes de conception du LMCE, tels que l'approfondissement de la modélisation des tâches d'accès à l'information ou l'utilisation de la dimension informationnelle des profils des utilisateurs pour la mise au point de cette indexation.

4.1.1.3. Articulation des modèles pour l'accès à l'information

Les trois composantes de la démarche d'accès à l'information que nous voulons mettre au point sont donc, comme illustré en Figure 22, les modèles de connaissance, le système d'information de l'organisation et l'utilisateur. Pour être générique, une telle démarche doit reposer sur trois modèles stables : les modèles du domaine d'activité de l'organisation, de son système d'information et des utilisateurs de ce système d'information.

Nous avons vu au chapitre 3 que le premier modèle est issu d'une phase de capitalisation des connaissances menée dans l'unité, à l'aide de la méthode MKSM. La problématique de modélisation du domaine d'activité de l'organisation ne se pose donc pas. Les méthodologies de conception des deux autres modèles, qui doivent être élaborés spécifiquement pour le LMCE, ont été développées aux paragraphes précédents.

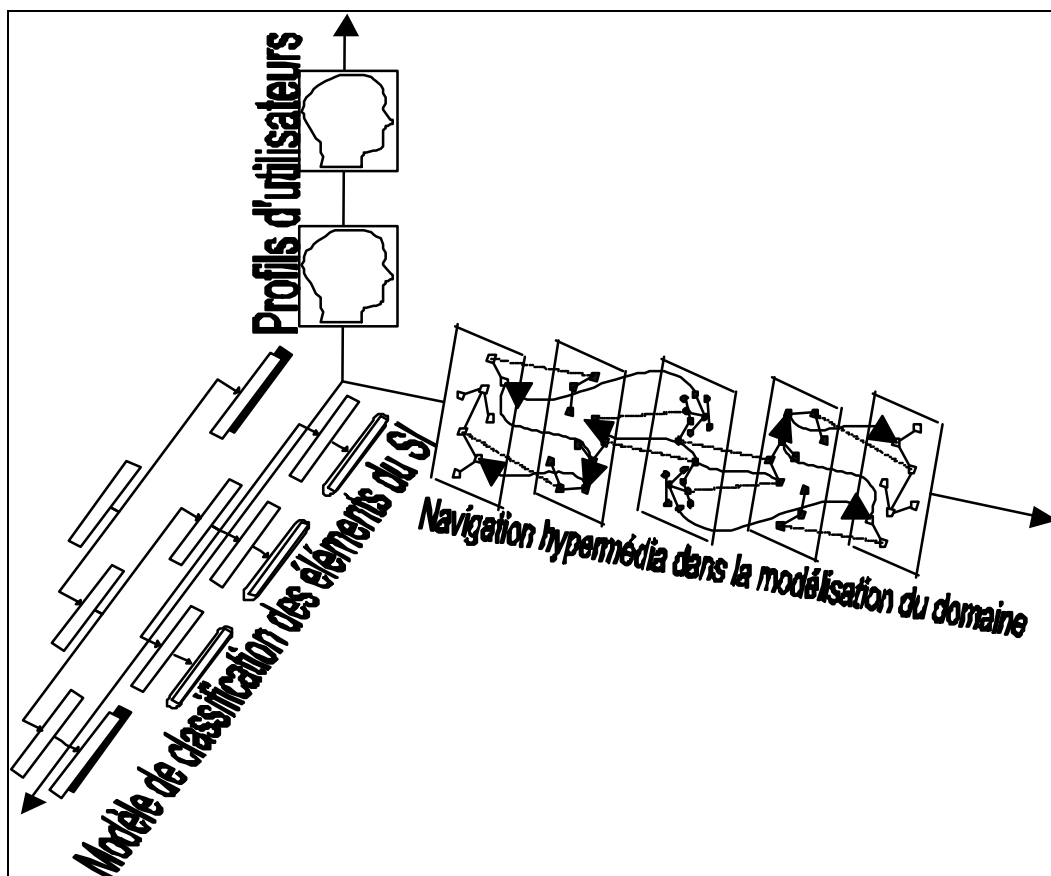


Figure 22. Les trois dimensions de notre démarche d'accès à l'information

Cette démarche d'accès à l'information comporte d'une part une stratégie d'accès à l'information spécifique et d'autre part un outil d'accès à l'information mettant en œuvre cette stratégie : le LMCE. La problématique de définition de cette stratégie et celle de la conception du LMCE sont développées aux paragraphes suivants.

4.1.2. *Vers une stratégie spécifique d'accès à l'information*

Nous avons vu au chapitre 2 les spécificités des stratégies d'accès à l'information analytiques et de parcours. Pour G. Marchionini, ces stratégies sont complémentaires et correspondent entre autres chez l'utilisateur à des problèmes d'information différents ([Marchionini 95]). Elles peuvent être combinées, par exemple pour restreindre le champ d'une requête documentaire à un sous-ensemble d'une collection de documents (voir chapitre 2).

C'est cette utilisation séquentielle des deux stratégies qui est présentée dans la version initiale du LCE de [Chaillot & Ermine 97]. L'utilisateur alterne navigation et interrogation : il suit tout d'abord une stratégie opportuniste de parcours de la collection de diagrammes, puis interroge la totalité du fonds sur un élément sélectionné dans le dernier diagramme exploré. Il

ne peut donc pas utiliser pour sa requête les éléments pertinents des diagrammes visualisés lors des étapes de navigation précédentes.

Or, les entretiens avec les experts nous apprennent que les tâches de recherche d'information sont souvent liées à des problèmes complexes associant plusieurs concepts. C'est d'ailleurs un des avantages de la méthode de modélisation MKSM que de pouvoir faire apparaître la complexité de ces relations sur les diagrammes de connaissances. La mise en relation de plusieurs éléments des diagrammes aboutit à une requête plus pertinente et plus utile que la simple exploration du fonds relatif à chaque concept. C'est pour répondre à cette nouvelle problématique que nous avons développé un autre type de stratégie d'accès à l'information.

Nous voulons ici mettre en place la possibilité d'interroger le fonds documentaire sur plusieurs concepts sélectionnés au cours de la navigation dans des diagrammes différents. Cela présente deux difficultés.

D'une part, il s'agit de mettre en œuvre de façon cohérente la recherche à partir d'un élément des diagrammes. En effet, tous les éléments n'ont pas la même « valeur sémantique » : une requête sur le contenu textuel d'un flux documentaire dans un diagramme d'activité n'est par exemple pas très pertinente, alors qu'elle l'est sur le contenu d'un concept ou d'un objet⁸¹. Cette difficulté peut par exemple être surmontée avec un typage des éléments des diagrammes et une réaction appropriée du système en fonction de chacun des types.

D'autre part, il s'agit d'articuler tous les éléments sélectionnés des diagrammes de façon interprétable à la fois par l'utilisateur et le SRI. Pour cela, la solution que nous allons explorer s'appuie sur des règles ergonomiques. Nous proposons, entre autres de permettre à l'utilisateur de procéder pas à pas en spécifiant explicitement la relation entre les concepts qu'il désire associer et en visualisant la requête qu'il est en train de construire.

4.1.3. *Problématique opérationnelle*

La problématique opérationnelle de cette thèse est d'étudier la conception d'un outil générique d'accès à l'IST combinant la navigation hypermédia dans des diagrammes de connaissance et l'interrogation d'un fonds documentaire. Cette étude fait donc appel aux disciplines de la conception logicielle et d'hypermédias et de l'ingénierie de l'information.

⁸¹ À titre d'exemple, dans le projet pilote sur la sûreté-criticité, une recherche sur le texte « Note d'étude de criticité » renvoie théoriquement toute une classe de documents. Par contre, la recherche sur le concept « Note d'étude de criticité » doit elle renvoyer des éléments qui ne sont pas des notes d'étude mais des informations méthodologiques sur la rédaction de ces documents.

En génie logiciel, des méthodes de modélisation sont utilisées avec plusieurs objectifs : l'analyse des besoins des utilisateurs finaux, la spécification des fonctionnalités et du comportement de l'interface et enfin la conception logicielle⁸². Les trois premières phases relèvent également de l'ergonomie. Une méthode de conception ergonomique s'appuie donc également sur une modélisation, qui peut être menée par ailleurs. Dans ce cas, cette double modélisation représente un surcoût pour les concepteurs et pour l'organisation où elle est menée⁸³. D'où l'intérêt d'utiliser une unique méthode de modélisation pour ces quatre objectifs. Nous faisons ici l'hypothèse que la méthode MKSM permet d'obtenir tous les éléments nécessaires à cette modélisation. Nous l'utilisons dans cette thèse pour les différentes étapes de la conception logicielle.

Nous avons vu au chapitre précédent que le LCE est un document hypermédia statique couplé à un système d'accès à l'information. Ces deux éléments possèdent un contenu prédéfini : la partie hypermédia est constituée du recueil de diagrammes de connaissances, figé par la méthode MKSM, et le système d'accès à l'information est un Moteur d'Indexation et de Recherche développé par ailleurs. La problématique de conception du LMCE est donc principalement ergonomique. Elle consiste à spécifier la mise en œuvre d'une interface pour ce système, à la fois utilisable et pertinente. En effet, une fois livré, le LMCE est « en compétition » avec les autres outils d'accès à l'information de l'organisation. Pour qu'il soit utilisé par les acteurs de cette organisation, il doit donc réaliser la fonction d'accès à l'information de façon plus pertinente que les autres outils. Cette pertinence tient à la fois à l'efficacité du SRI utilisé, à la démarche d'accès à l'information et à la qualité ergonomique de l'interface.

D'autre part, dans les cas extrêmes, c'est-à-dire pour les utilisateurs capables de formaliser directement et rapidement leur problème d'information, le LMCE doit pouvoir être utilisé comme un SRI à base d'interrogation classique.

C'est pourquoi nous abordons cette problématique de conception non par l'aspect fonctionnel du LMCE – a priori, cet outil ne fait qu'associer d'une façon particulière un navigateur W3 et un moteur de recherche – mais par son aspect IHM. L'essentiel de la problématique opérationnelle de cette thèse se situe donc au niveau de la conception de l'interface du LMCE,

⁸² Par exemple, une méthode de conception objet comme OMT permet de définir non seulement les fonctionnalités du système à réaliser, mais aussi son architecture logicielle. Dans certains cas simples, elle permet même de générer le code correspondant.

⁸³ Du fait du temps de travail des individus qu'elle mobilise.

dont le comportement régit l'utilisation des fonctionnalités proposées par ces outils logiciels complexes.

Le but de notre étude est donc de spécifier, réaliser et tester un tel outil. Les spécifications se doivent d'être exhaustives pour permettre l'adaptabilité de l'outil aux différents projets et automatisables pour permettre la génération grâce à un atelier de production (l'objectif est l'industrialisation du LMCE). Ces spécifications sont réalisées en deux temps.

À un niveau générique, elles permettent la mise au point d'une architecture conceptuelle. La problématique de cette étape est notamment d'évaluer la part de chaque modèle dans le processus de conception et de déduire les différents composants nécessaires au fonctionnement du LMCE.

À un niveau technique, il s'agit de proposer une implémentation de l'architecture conceptuelle du LMCE. Les contraintes technologiques doivent être prises en compte pour aboutir à une description de l'architecture logicielle du LMCE. Dans cette description, une distinction doit être faite entre les composants réutilisables et ceux issus de la phase de modélisation.

4.1.4. De la pertinence relative de la problématique de l'accès à l'information et de l'approche de la gestion des connaissances

À partir d'une modélisation multi-points-de-vue riche mais complexe, nous définissons donc trois modèles interreliés qui vont permettre d'aider un individu identifié dans une situation de travail donnée à formuler son besoin d'information. En tant qu'utilisateur du LMCE, ce dernier visualise et interagit avec ces relations par le comportement de son interface.

Il se voit par exemple, du fait de son profil (i.e. à cause de sa fonction ou de ses compétences), limiter l'accès à certaines sources d'informations. Ceci illustre la relation qui existe entre le modèle de l'utilisateur et celui du SI. De même, le lien entre le profil de l'utilisateur et le modèle du domaine se manifeste lors de la sélection des collections de diagrammes qui lui sont proposés en fonction de son activité.

La relation entre le modèle du domaine (i.e. les diagrammes de connaissances) et celui du SI est plus complexe : elle nécessite la mise en correspondance des éléments des diagrammes et du système d'information. C'est également cette relation qui est au cœur de la problématique de facilitation de formalisation du problème d'information de l'utilisateur. Pour celui-ci, cette relation est matérialisée par la stratégie d'accès à l'information que lui propose l'interface du LMCE.

La méthodologie de conception d'un LMCE permet d'obtenir chacun de ces modèles, parfois les uns en fonction des autres. La démarche de spécification ergonomique employée doit normalement permettre une bonne utilisabilité de cet outil. Toutefois, il est intéressant de s'interroger a posteriori sur le statut de ces relations, à la fois par rapport à l'outil LMCE et à l'utilisateur de cet outil.

Cette réflexion est présentée au chapitre 5. Elle s'appuie sur une vision fonctionnelle qui permet de mettre en lumière les interactions entre ces disciplines. L'objectif est d'évaluer la pertinence de l'approche spécifique d'accès à l'information par recontextualisation définie dans ce chapitre. De cette réflexion émerge alors une discussion sur la pertinence de l'approche de la gestion des connaissances pour la résolution des problèmes d'information qui se posent aux individus dans le cadre de l'accomplissement d'une activité à dominante scientifique et technique.

L'utilisation de la méthode de gestion des connaissances MKSM comme cadre formel pour la conception, la spécification et l'analyse a posteriori de l'outil est également discutée. Il s'agit d'introduire une réflexion plus théorique sur le statut des modèles MKSM et des paradigmes systémique et sémiotique sur lesquels cette méthode se fonde par rapport à la méthode de modélisation nécessaire pour mettre au point un LMCE.

Un bilan prospectif permet de synthétiser les conclusions tirées de la discussion des différents aspects de cette approche. Ce bilan nous amène à souligner les avancées et les limitations du travail réalisé et à envisager des perspectives sur de futures évolutions du système et de l'approche.

Cette discussion méthodologique doit donc s'appuyer sur une architecture détaillée du LMCE. L'étude de conception aboutissant à cette architecture est présentée dans la partie suivante.

4.2. Démarche de conception du LMCE

Jean-Louis Ermine, fondateur de la méthode MKSM, avait coutume de présenter le LCH (i.e. la mise au format hypermédia d'un livre de connaissances) comme « *un programme qui ne fait rien* ». Dans les versions qui ont été réalisées précédemment (voir chapitre 3), le « programme » LCH est en fait un navigateur qui permet de visualiser le document hypermédia constitué par une collection de diagrammes. Il est donc normal que le LCH, en tant que document hypermédia statique ne fasse rien. En particulier, la mise au format HTML de cette collection permet l'utilisation d'un navigateur existant, et le « programme » est alors décorrélé de l'hyperdocument LCH.

Le LMCE est lui aussi un hyperdocument, puisqu'il est destiné à être visualisé à l'aide d'un navigateur W3. Cependant, il s'agit désormais d'un hyperdocument dynamique, puisqu'il permet entre autres la gestion d'une interface personnalisée et l'accès à l'information. À ce titre, ce n'est plus un « programme » au sens autonome du terme, mais en contrepartie, il fait désormais quelque chose.

Dans cette partie, nous présentons le travail de conception de cet hyperdocument dynamique réalisé pendant cette thèse. Comme nous l'avons déjà souligné, ce travail s'appuie sur une démarche de conception ergonomique qui a pour but de faciliter l'utilisabilité du LMCE. Cette démarche de conception s'appuie elle-même sur trois modèles et deux outils conceptuels qui permettent de définir les fonctionnalités du LMCE et qui sont présentés respectivement dans une première et une deuxième sous-parties. La mise au point de l'interface fait l'objet de la troisième sous-partie. Les spécifications et l'architecture de l'outil sont alors proposées dans une quatrième sous-partie. Dans la perspective de l'industrialisation du processus de conception du LMCE, nous étudions le cycle de vie de cet outil en fonction des unités où il a à être installé.

Le travail de conception mentionné ici décrit l'approche et propose les spécifications techniques d'un prototype de LMCE. Actuellement, seul le prototype de LCE réalisé par Jean-Marc Charlot à partir de la modélisation réalisée lors du projet-pilote est totalement implémenté. Ce prototype est décrit dans cette partie, à la fois pour permettre d'évaluer les développements réalisés et pour positionner le LMCE par rapport à l'existant. Nous indiquons notre contribution à ce prototype. Ce prototype ne correspond cependant pas à la description du LMCE faite dans cette thèse.

Suite à la livraison tardive de l'atelier MNESIK, le développement d'un autre prototype, plus proche de l'outil décrit ici, est toujours en cours (finalement sans l'aide de cet atelier). Les aspects techniques mentionnés ci-dessous décrivent ce prototype partiellement implémenté.

4.2.1. Les trois dimensions de l'expression du besoin d'information avec le LMCE

Dans ce paragraphe, nous présentons les descriptions des trois modèles conceptuels du domaine, des utilisateurs et de l'information sur lesquels s'appuie la conception du LMCE. Ces descriptions sont abordées sous des aspects théoriques puis techniques. La description technique indique l'approche qui a conduit au choix d'une structure interne du modèle et décrit cette structure. La description technique illustre l'implémentation de cette structure.

4.2.1.1. Le référentiel diagrammatique du domaine d'activité

Le domaine d'activité de l'organisation est le point de départ de notre approche d'accès à l'information. Il se fonde sur une collection de diagrammes de connaissances obtenue grâce à la méthode MKSM. Nous cherchons ici à définir l'ensemble du modèle qui sert de base pour notre démarche d'accès à l'information. Grâce aux diagrammes de connaissances, ce modèle contient l'ensemble des informations concernant le domaine d'activité dans le contexte duquel l'information est accédée. C'est pourquoi il est qualifié de référentiel diagrammatique de ce domaine.

4.2.1.1.1. Aspects théoriques

Le modèle du domaine contient l'ensemble des diagrammes de connaissances issus de la modélisation de l'activité de l'organisation. Cette modélisation est présentée dans le LMCE sous forme de document hypermédia. Elle permet à la fois à l'utilisateur de naviguer dans l'activité de l'organisation et d'accéder à l'information. Nous allons donc examiner cette collection de diagrammes en fonction de ce double objectif. Le but de ce paragraphe est d'aboutir à la définition d'une structure interne de ce modèle.

4.2.1.1.1.1. Structuration du domaine

Nous avons vu au chapitre 3 que la méthode MKSM permet d'articuler l'ensemble des diagrammes de connaissances décrivant un domaine. Cette articulation tient compte des différents points de vue représentés par chacun des formalismes diagrammatiques. Cependant, la méthode MKSM a été conçue dans le but de capitaliser le savoir-faire d'un expert, par le

biais de ces diagrammes et des nombreuses relations qui existent entre eux, et de restituer ce savoir-faire à n'importe quel type de lecteur du livre de connaissances.

Il apparaît alors que ce type de rédaction uniquement centré sur l'expert, peut parfois, malgré tout l'intérêt porté à la compréhensibilité des modèles, ne pas totalement satisfaire le lecteur, en lui fournissant soit trop, soit trop peu de détails, en fonction de son niveau d'expertise ou du « niveau de granularité » avec lequel il s'intéresse au domaine. Par ailleurs, il arrive, au cours d'un projet de capitalisation des connaissances dans une organisation, de devoir s'intéresser à plusieurs types d'experts, travaillant dans le même contexte, mais dans des domaines d'expertise différents. On ne peut alors réaliser un livre de connaissances pour chacun d'entre eux. Cela empêcherait de définir les liens entre ces sous-parties de l'organisation, ce qui revient à abandonner l'approche systémique sur laquelle est bâtie MKSM. Il pourrait donc être intéressant de tenir compte de tous ces sous-systèmes de l'organisation dans la modélisation globale de celle-ci, quitte à s'éloigner quelque peu de notre paradigme sémiotique de départ. En contrepartie, on conserve le paradigme systémique en proposant une description unique de l'organisation à plusieurs niveaux de décomposition.

La modélisation obtenue comporte alors plusieurs entités, baptisées *sous-domaines*. En fonction de leurs spécificités, ceux-ci peuvent être totalement disjoints, partager certains pans de la modélisation, voire la quasi-totalité de celle-ci⁸⁴.

La collection de diagrammes de connaissances plus l'ensemble de relations les unissant deviennent alors un « système de diagrammes de connaissances » (au sens de Le Moigne). C'est la mise au format hypermédia de ce système qui constitue alors le LCH. À chaque sous-domaine identifié correspond un point d'entrée dans le LCH.

4.2.1.1.1.2. Structuration des sous-domaines

Chaque sous-domaine regroupe une collection de diagrammes issus d'une partie de la modélisation des connaissances. À ce titre, un sous-domaine de la modélisation possède une structure interne et les diagrammes de connaissances sont articulés comme lors d'une modélisation centrée sur un expert. Les liens vers des diagrammes d'autres sous-domaines sont référencés en tant que tels dans les diagrammes et traités comme des liens statiques externes vers d'autres types de sources d'information (i.e. des documents, voir chapitre 3).

⁸⁴ Par exemple, dans le projet opérationnel sur la maîtrise du risque de criticité, six sous-domaines ont été identifiés. Pour chacun d'eux, une phase d'interviews/modélisation a été menée et a débouché sur la rédaction d'une collection complète de diagrammes décrivant chaque type d'activité, à l'exception du modèle du domaine, qui décrit les processus physiques régissant la criticité nucléaire, communs à tous ces sous-domaines.

Nous n'avons cependant pas défini une structure hiérarchique rigide pour chaque sous-domaine. En effet, certains diagrammes peuvent être communs à plusieurs sous-domaines. La duplication de ces diagrammes pour chacun d'eux induirait, en plus d'un « gaspillage » d'espace-disque, la perte d'une relation importante entre ces sous-domaines. C'est pourquoi nous avons choisi, plutôt que d'isoler ces diagrammes, de référencer pour chacun d'eux le(s) domaine(s) au(x)quel(s) ils appartiennent.

4.2.1.1.1.3. Structuration des diagrammes

Le LCH doit être porteur de sens pour l'utilisateur qui y navigue. Le caractère compréhensible du LCH est déterminé à la fois par le fonds (le contenu sémantique des diagrammes et des relations) et la forme. La qualité du fonds découle d'une part de celle de la modélisation, et d'autre part de celle de la méthode MKSM. Ni l'une ni l'autre ne sont discutées ici. La mise en forme du LCH, et notamment la facilitation de la navigation via les cartes conceptuelles, est présentée dans la partie concernant les composants d'interface du LMCE.

4.2.1.1.1.4. Structuration des éléments des diagrammes

La collection de diagrammes est utilisée dans le LMCE, non seulement pour la navigation hypermédia, mais aussi pour l'accès à l'information. Dans cette optique, la définition du modèle du domaine tient compte de cette caractéristique du LMCE.

Un diagramme est constitué d'arcs et de **nœuds**. Les arcs représentent les relations sémantiques intra et inter diagrammes définies dans la méthode MKSM⁸⁵. Les **nœuds** peuvent être représentés par un ensemble de champs, indiquant leur position, type, texte... L'accès à l'information se fait grâce aux **nœuds**.

Nous avons vu précédemment que les champs type et texte des **nœuds** sont tous deux importants pour la recherche d'information. Cependant, si le filtrage peut se faire en fonction du type des **nœuds** (il existe un nombre fini de types), des essais préalables de recherche d'information sur le contenu textuel exact d'un **nœud** se sont avérée peu convaincants. En effet, la méthode MKSM permet et recommande d'explicitier chacun des **nœuds** des diagrammes. Tout en restant synthétiques, certains **nœuds** peuvent donc contenir plus d'une dizaine de mots, qui constituent une chaîne de caractères introuvable pour le moteur de

⁸⁵ Ces relations sont détaillées dans [Médini 97]. Les relations inter-diagrammes sont matérialisées par les liens hypermédiés dans le LCH.

recherche utilisé. À plus forte raison, une requête constituée à partir du croisement des textes de plusieurs nœuds a peu de chances de donner des réponses⁸⁶.

C'est pourquoi nous avons choisi de décrire (ou de reformuler) le texte de chaque nœud dans un champ supplémentaire. Cette reformulation contient une liste des descripteurs du texte, rédigée en fonction de la syntaxe du moteur de recherche utilisé. Par exemple, une activité dont le texte est « Conception d'une installation mettant en œuvre des matières fissiles » est décrite, dans le langage VDK utilisé par le moteur de recherche TOPIC d'Acrobat Search, par « conception NEAR (installation OR INB) AND "matière fissile" »⁸⁷.

Ces champs de reformulation sont alors à renseigner par le chargé de projet lors de la rédaction du diagramme. Le renseignement de ces champs peut faire l'objet d'une discussion avec l'expert. Dans le prototype de [Charlot & al. 00], ces champs sont automatiquement renseignés à la génération du document hypermédia dans le cas où ils ne l'ont pas été lors de la rédaction des diagrammes. Dans ce cas, leurs contenus sont ceux des champs texte des éléments correspondants.

4.2.1.1.2. *Aspects techniques*

Le modèle du domaine, comme les deux autres modèles, sont stockés dans des bases XML. À chaque diagramme de connaissances correspond un fichier XML. Chaque fichier XML est associé à une DTD qui permet, au besoin, de vérifier la syntaxe du fichier. Il n'est pas dans nos objectifs de mettre au point une DTD pour chaque type d'élément des formalismes MKSM. Ce type d'exercice a déjà été réalisé auparavant, notamment pour l'opérationnalisation de la méthode (voir chapitre 3) et ne présente pas d'intérêt pour la compréhension de la suite. Nous supposons donc ici les documents XML valides et bien formés, puisqu'ils sont obtenus automatiquement par traitement des fichiers sources des diagrammes. Ici, la DTD ne sert donc qu'à définir les références utilisées dans les diagrammes (notamment les valeurs numériques associées aux sous-domaines⁸⁸). Le début et la fin d'un fichier XML du modèle du domaine sont donc les suivants :

⁸⁶ Cela dépend bien entendu du moteur de recherche utilisé. Les essais ont été réalisés sur un échantillon du corpus de documents en pdf du projet-pilote. Cet échantillon a été indexé à l'aide du logiciel Acrobat Catalog® et la recherche d'information menée avec Acrobat Search®. Peut-être un moteur de recherche utilisant des dictionnaires de reformulation et des traitements linguistiques (comme SPIRIT®) aurait-il donné de meilleurs résultats. Cependant, SPIRIT® n'indexait pas les documents pdf à l'époque.

⁸⁷ Ce moteur de recherche utilisant la troncature des mots, il n'est pas nécessaire de différencier "matière fissile" de "matières fissiles".

⁸⁸ Les sous-domaines sont traités par références pour permettre leur réutilisation dans les profils des utilisateurs sans modification de la structure générique de ces profils.


```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE diagramme SYSTEM "diagramme.dtd">
<DIAGRAMME>

...

</DIAGRAMME>
```

La structure des documents XML est constituée de deux types de champs : les champs d'en-tête du fichier HTML correspondant et les éléments. Les champs d'en-tête délimités, par une balise « HEAD », permettent de stocker les informations sur le diagramme : titre, auteur, organisation propriétaire du diagramme, formalisme et sous-domaine. Ce type d'en-tête, hérité des documents HTML, est conservé afin de faciliter le parsing en HTML et de rendre les fichiers XML plus « parlants ». L'élément HEAD du fichier XML correspondant au diagramme d'activité du projet-pilote donné au chapitre 3 est présenté ci-dessous.

```
<HEAD>
  <TITLE>Activités liées au cycle de vie d'une INB</TITLE>
  <REDACTEUR>Lionel Médini</REDACTEUR>
  <PROPRIETAIRE>IPSN/DPEA/SEC</PROPRIETAIRE>
  <FORMALISME>Activite</FORMALISME>
  <SOUSDOMAINES>
    &INB;
  </SOUSDOMAINES>
</HEAD>
```

Les champs d'éléments sont délimités par une balise « MAP ». En effet, le logiciel graphique (Visio) que nous avons utilisé pour rédiger les diagrammes permet d'exporter les diagrammes au format HTML en générant un fichier image (GIF ou JPEG, au choix) et en « mappant » les zones cliquables sur cette image⁸⁹. Le fichier HTML produit contient donc des balises « IMG » et « MAP » que nous avons reproduites à l'identique dans les fichiers XML.

```
<IMG SRC="INB-Activite2.jpg" ALT="Page-2" BORDER="0" USEMAP="#visImageMap">
</IMG>
<MAP NAME="visImageMap">

...

</MAP>
```

Nous avons réalisé deux programmes qui exploitent la fonctionnalité d'exportation en HTML de Visio®. Le premier est un pré-traitement qui renseigne les champs de reformulation de tous les nœuds des diagrammes dans le fichier source. Pour cela, nous avons posé des liens hypermédia sur ces nœuds et utilisé les champs référence de ces liens pour passer le texte reformulé du format propriétaire de Visio® en HTML. Le second est un post-traitement qui « traduit » les fichiers HTML générés en XML (en refermant les balises non fermées et en

transformant en unicode tous les caractères spéciaux du texte des diagrammes), afin d'obtenir du XML bien formé. Ces deux programmes sont des macros écrites en Visual Basic pour Applications, et sont utilisables, pour le premier sous Visio et pour le second sous Word. Nous récupérons ainsi un fichier XML bien formé qui « mappe » une image de diagramme de connaissances. Chacun des nœuds contient donc un lien dont le champ référence contient un ou plusieurs liens existant dans le fichier original s'il en existait et le texte reformulé destiné à la recherche d'information. Ce format de fichiers correspond au format d'entrée de l'atelier Arelice de Jean-Marc Charlot qui permet de mettre au point, en Java, les fichiers HTML dynamiques définitifs utilisés pour le démonstrateur⁹⁰.

Dans le présent prototype, nous proposons une structuration différente des fragments correspondant aux nœuds des diagrammes. Le contenu de ces nœuds reste proche de celui du prototype de Jean-Marc Charlot, mais nous y avons explicité le contenu des champs. Par ailleurs, cette syntaxe permet de faire correspondre à chaque élément autant de liens qu'il y en a dans le fichier original. Des éléments spécifiques à la recherche d'information sont également disponibles. Ces éléments se composent de plusieurs fragments spécifiques aux moteurs de recherches identifiés avec le prototype. Chacun de ces fragments permet de définir un texte et une valeur de reformulation différente du texte du nœud correspondant au moteur de recherche considéré⁹¹. Un élément d'un diagramme se présente alors sous la forme suivante :

```
<NOEUD>
  <ID>activitedecomposee3</ID>
  <TYPE>activitedecomposee</TYPE>
  <TEXTE>Conception d'une installation mettant en oeuvre des matières
fissiles</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>154,189,267,189,267,121,154,121,154,189</COORDS>
  <LIEN>
    <TEXTE>Détailier l'activité</TEXTE>
    <DEST>INB-activite2.html</DEST>
  </LIEN>
  <LIEN>
    <TEXTE>Liste des matières fissiles : modèle des concepts</TEXTE>
    <DEST>INB-concept4.html</DEST>
```

⁸⁹ Avec plus ou moins de bonheur : il arrive assez rarement que la zone cliquable définie dans le fichier HTML généré soit située à proximité de la zone à cliquer de l'image. Nous avons d'ailleurs dû réaliser un programme de « remapping » visuel de ces zones.

⁹⁰ Cet atelier génère, entre autres, les cartes conceptuelles définies dans la méthode MKSM.

⁹¹ Les moteurs de recherche que nous avons identifiés ici sont Acrobat Search® avec lequel nous avons effectué l'indexation du fonds et SPIRIT® avec lequel nous espérons à terme pouvoir réaliser cette indexation. Ce dernier outil fonctionnant en langage naturel et disposant de ses propres dictionnaires de reformulation, l'une des valeurs de reformulation correspondante est vide, ce qui cause l'interrogation en fonction du texte du nœud. Les éléments de type « interrogation contextuelle » sont destinés à accélérer la recherche. Ils sont définis au paragraphe 5.1.2.4.1.

```

</LIEN>
<RECHERCHE>
  <MIR>Acrobat Search</MIR>
  <TEXTE>Interrogation contextuelle sur cette activité</TEXTE>
  <REFO>"note d'étude de criticite" AND "dossier de conception" AND "rapport
preliminaire de surete"</REFO>
</RECHERCHE>
<RECHERCHE>
  <MIR>Acrobat Search</MIR>
  <TEXTE>Interrogation sur le texte de cette activité</TEXTE>
  <REFO>conception NEAR (installation OR INB) AND "matiere fissile"</REFO>
</RECHERCHE>
<RECHERCHE>
  <MIR>SPIRIT</MIR>
  <TEXTE>Interrogation contextuelle sur cette activité</TEXTE>
  <REFO>note d'étude de criticité, dossier de conception, rapport
préliminaire de sûreté</REFO>
</RECHERCHE>
<RECHERCHE>
  <MIR>SPIRIT</MIR>
  <TEXTE>Interroger le fonds documentaire</TEXTE>
  <REFO/>
</RECHERCHE>
</NOEUD>

```

Au final, nous obtenons un ensemble de fichiers à la fois « lisibles » et utilisables pour la génération du LCH. La visualisation de ces diagrammes est régie par des feuilles de style XSL, qui sont décrites plus loin. Ces documents XML permettent également la génération de cartes conceptuelles telles que des arbres fonctionnels, dictionnaires des acteurs ou des concepts, définis par la méthode MKSM et présents dans le prototype de Jean-Marc Charlot.

4.2.1.2. Les profils d'utilisateurs

Par définition, une architecture de DVP nécessite un modèle de l'utilisateur capable de fournir les informations à partir desquelles l'interface sera personnalisée. Nous avons vu au chapitre 2 l'intérêt de disposer d'un modèle de l'utilisateur qui tienne compte de ses caractéristiques autant pour la navigation dans les diagrammes de connaissances que pour l'accès à l'information. Nous avons ensuite, au chapitre 3, identifié dans la littérature trois catégories de caractéristiques pertinentes pour ces modèles. Nous fixons ici le détail de ces caractéristiques.

Le niveau d'expertise du domaine. L'utilisateur peut être novice, apprenant ou expert dans chacun des sous-domaines modélisés.

Les préférences informationnelles. L'utilisateur peut sélectionner une stratégie d'accès à l'information privilégiée. Les stratégies proposées sont la stratégie spécifique d'accès à l'information par les diagrammes de connaissances, une interrogation directe de la base par la saisie d'une requête textuelle ou une interrogation de la base par des champs de description

des documents (titre, auteur, date, langue...). L'utilisateur peut aussi choisir s'il préfère un filtrage « faible » ou « important » de sa recherche d'information. Il indique enfin son niveau de connaissance des outils d'accès à l'information. Il s'agit de déterminer de façon booléenne s'il est capable d'utiliser le langage de requêtes du moteur de recherche auquel est relié le LMCE.

Les préférences de visualisation. Ces préférences permettent de déterminer les types d'informations présents à l'écran, leur nombre et leur disposition. Pour le premier élément, nous avons choisi d'affecter à chacun des composants de l'interface une valeur entière permettant à l'utilisateur de spécifier l'intérêt que cet élément présente pour lui. Il pourra donc déterminer s'il veut que cet élément ne soit jamais présent à l'écran, qu'il soit présent quand cela est jugé nécessaire par le système ou qu'il soit toujours présent. Les réponses non, peut-être ou oui correspondent respectivement aux valeurs 0, 1 ou 2. Le nombre d'éléments maximum présents à l'écran détermine le nombre de cadres HTML à afficher. Leurs dispositions comprennent la taille (horizontale et verticale) et les coordonnées du coin supérieur gauche. Ces quantités sont données en pourcentage par rapport à la taille de l'écran. Elles sont modifiées dynamiquement lors de tout changement de ces dispositions de la part de l'utilisateur. Ces préférences doivent bien entendu être cohérentes entre elles. La validation de celles-ci est donc nécessaire.

4.2.1.2.1. *Aspects techniques*

Comme pour les deux autres modèles, le modèle des utilisateurs est constitué d'une base de documents XML. N'étant pas destinés à être visualisés, ces documents ne comportent que des informations « utiles » et aucune en-tête comme c'est le cas pour les diagrammes de connaissances. Le fichier XML commence donc par les déclarations d'usage et l'identification de l'utilisateur.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE diagramme SYSTEM "diagramme.dtd">
<PROFIL>

  <UTILISATEUR>
    <NOM>Médini</NOM>
    <PRENOM>Lionel</PRENOM>
    <PASSWORD>toto</PASSWORD>
  </UTILISATEUR>
  ...
</PROFIL>
```

L'utilisateur est repéré par un nom et un mot de passe. Nous ne discutons pas ici la pertinence ni la mise en œuvre des aspects de sécurité du prototype. La connexion par mot de passe est

uniquement destinée à faire comprendre à l'utilisateur que ses préférences ne sont pas directement accessibles à d'autres acteurs de l'organisation.

Les niveaux d'expertise de l'utilisateur sont ensuite définis en fonction des sous-domaines. Ici, nous avons défini l'utilisateur Medini comme expert dans les sous domaines de conception des INB et des expériences critiques, comme apprenant dans les sous-domaines des transports de matières fissiles et des codes de calculs et comme novice dans les deux autres sous-domaines⁹². Les noms des sous-domaines et les niveaux d'expertise sont des références à des entiers (0 à 5 pour les sous-domaines et 0 à 2 pour les niveaux d'expertise) définis dans la DTD.

```
<EXPERTISE>
  <SOUS-DOMAINES>
    <NOM>&INB; </NOM>
    <NIVEAU>&expert; </NIVEAU>
  </SOUS-DOMAINES>
  <SOUS-DOMAINES>
    <NOM>&Transports; </NOM>
    <NIVEAU>&expert; </NIVEAU>
  </SOUS-DOMAINES>
  <SOUS-DOMAINES>
    <NOM>&Experiences; </NOM>
    <NIVEAU>&apprenant; </NIVEAU>
  </SOUS-DOMAINES>
  <SOUS-DOMAINES>
    <NOM>&Codes; </NOM>
    <NIVEAU>&apprenant; </NIVEAU>
  </SOUS-DOMAINES>
  <SOUS-DOMAINES>
    <NOM>&Standards; </NOM>
    <NIVEAU>&novice; </NIVEAU>
  </SOUS-DOMAINES>
  <SOUS-DOMAINES>
    <NOM>&Diffusion; </NOM>
    <NIVEAU>&novice; </NIVEAU>
  </SOUS-DOMAINES>
</EXPERTISE>
```

Les trois types de préférences informationnelles sont ensuite mentionnés. La stratégie d'accès à l'information peut être cumulative (voir paragraphe suivant), par champs ou directe. Le filtrage peut être défini comme large ou restreint. Le langage d'interrogation du moteur d'indexation et de recherche est connu ou inconnu. La définition de ces caractéristiques utilise des références à des valeurs numériques (0 à 2 pour la stratégie d'accès à l'information et 0 ou 1 pour le filtrage et le langage d'interrogation) définies dans la DTD.

```
<INFORMATION>
```

⁹² En fait, l'utilisateur en question est loin de posséder les niveaux d'expertise dans ces sous-domaines. Cependant, en tant que concepteur du système, nous nous sommes octroyé les niveaux d'expertises les plus élevés dans les sous-domaines où la modélisation est la plus complète, de façon à avoir accès à toutes les fonctionnalités du système.

```

<STRATEGIE>&cumulative;</STRATEGIE>
<FILTRAGE>&large;</FILTRAGE>
<LANGMIR>&connu;</LANGMIR>
</INFORMATION>

```

Enfin, les préférences de visualisation sont définies. Chaque élément de l'interface est défini par un souhait de visualisation et les caractéristiques d'affichage du cadre correspondant.

```

<VISUALISATION>
  <CADRES>
    <DIAGRAMME>
      <SHOW>&OUI;</SHOW>
      <WIDTH>80%</WIDTH>
      <HEIGHT>80%</HEIGHT>
      <XPOS>20%</XPOS>
      <YPOS>10%</YPOS>
    </DIAGRAMME>
    ...
  </CADRES>
</VISUALISATION>

```

Les valeurs oui, peut-être et non (resp. 0, 1 et 2) de l'élément SHOW sont définies dans la DTD. La validation de la cohérence de ces données est faite dans la DTD. Il s'agit de vérifier qu'aucun cadre n'est plus grand que l'écran et que les cadres affichés ne se chevauchent pas.

Un exemple de fichier XML de profil complet est donné en annexe.

4.2.1.2.2. *Profils par défaut et actualisation des profils*

La modélisation de l'activité de l'organisation est effectuée auprès de quelques experts. Elle ne peut donc tenir compte des particularités de tous les individus amenés à utiliser le LMCE. Toutefois, il nous est possible de définir un ensemble de *profils d'utilisation*, qui serviront de bases aux futurs utilisateurs pour définir leur profil personnalisé. Nous définissons ici les profils d'utilisation comme la réunion des paramètres d'expertise et d'accès à l'information des profils présentés plus haut. Les préférences de visualisation sont discutées ensuite.

4.2.1.2.2.1. **Les profils d'utilisation**

Nous faisons ici l'hypothèse que les niveaux d'expertise du domaine et des outils d'accès à l'information influent sur le niveau de perception du problème d'information. Cette hypothèse – développée au paragraphe suivant – n'est pas très éloignée de celle de [Lainé-Cruzel & al. 96], pour qui ce niveau de perception influe sur la précision de la requête. En admettant qu'il sache manipuler le langage de requêtes de l'outil, un expert dans un sous-domaine formulera donc des requêtes dans ce sous-domaine plus précises qu'un novice ou un apprenant. À ce titre, et malgré le nombre important de facteurs qui entrent en ligne de compte pour la définition d'une stratégie d'accès à l'information, nous pouvons lui associer une stratégie

d'accès à l'information par interrogation. Cet accès à l'information peut être direct ou par champs.

Pour sélectionner l'une de ces deux stratégies, nous introduisons ici une distinction entre les domaines modélisés. Dans le projet-pilote, nous avons distingué deux types de sous-domaines de la sûreté-criticité nucléaire : les sous-domaines d'ingénierie et de recherche. D'après les experts interviewés dans chacun de ces sous-domaines, il ressort que les types d'informations recherchés sont différents. Un ingénieur a par exemple besoin de notes précises sur des études existantes dont il a connaissance. Dans ce cas, il est capable d'en indiquer l'auteur, la date et parfois même le titre exact, sinon une série de descripteurs précis. Pour ce type de recherche, il a donc recours à une interrogation par champs. Un chercheur s'intéresse lui à des articles ou des ouvrages traitant d'un sujet précis mais sans en connaître préalablement l'existence. Il préfère donc pouvoir formuler librement sa requête par interrogation directe.

Cette distinction entre ingénieur et chercheur se retrouve également dans le type de filtrage mais est ici indépendante de son niveau de connaissance de l'outil. Il ressort des entretiens avec les experts qu'un ingénieur utilise un filtrage plus restreint qu'un chercheur. Cette règle est valable quel que soit le niveau d'expertise de l'utilisateur dans le sous-domaine considéré.

Lorsqu'il ne connaît pas le langage d'interrogation du moteur de recherche utilisé, nous formulons l'hypothèse que la stratégie d'accès à l'information cumulative peut aider un expert d'un sous-domaine de recherche à formaliser son besoin d'information. Cette hypothèse est développée au paragraphe suivant. Cette hypothèse n'est pas faite pour les sous-domaines d'ingénierie, car l'interrogation par champs ne requiert pas l'utilisation de ce langage d'interrogation.

Les utilisateurs apprenants n'ont pas par définition un modèle clair de tous les concepts du domaine. La présentation de ce domaine à l'écran doit a priori les aider à formuler leur requête. C'est pourquoi nous leur associons la stratégie cumulative de construction de la requête par la navigation dans les diagrammes de connaissances (voir paragraphe suivant).

D'après les experts interviewés, la présentation à des individus novices des éléments d'information disponibles dans l'organisation n'est pas souhaitable. Cette information ne peut en effet qu'être incomprise ou mal interprétée⁹³. Or, la mécompréhension d'une information dans le domaine de la sûreté-criticité nucléaire peut avoir des conséquences fâcheuses (comme la diffusion d'une interprétation erronée de cette information). Les utilisateurs

⁹³ Pour avoir eu accès à ces informations et en tant que novice dans le domaine, nous confirmons ce jugement.

novices n'effectuent donc pas de recherche d'information dans les sous-domaines concernés. Seul l'accès à certains éléments d'information leur est autorisé. Ces éléments d'information sont alors reliés aux diagrammes par des liens statiques et ont le même statut que les diagrammes de connaissances (voir chapitre 3).

Le tableau suivant récapitule les caractéristiques des différents profils d'utilisation.

Profil d'utilisation	Stratégie d'accès à l'information	Type de filtrage
Expert ingénieur	Par champs	Restreint
Expert chercheur connaissant le langage du MIR	Par interrogation directe	Large
Expert chercheur ne connaissant pas le langage du MIR	Cumulative	Large
Apprenant ingénieur	Cumulative	Restreint
Apprenant chercheur	Cumulative	Large
Novice	Navigation seule	Néant

Tableau 4. Les caractéristiques des différents profils d'utilisation.

Chaque utilisateur « déclare » au système ses niveaux d'expertise dans chaque sous-domaine et sa connaissance ou non du langage d'interrogation. Le profil d'utilisation retenu sera celui correspondant à son niveau d'expertise le plus élevé. Dans le cas où l'utilisateur possède des niveaux d'expertise équivalents dans des domaines d'ingénierie et de recherche, la stratégie d'accès à l'information et le type de filtrage sont laissés au choix de l'utilisateur. Celui-ci peut d'ailleurs modifier ces préférences.

4.2.1.2.2.2. Les préférences de visualisation

Les préférences de visualisation par défaut sont celles présentées plus haut. Ces préférences peuvent être modifiées en changeant la taille des cadres dans le navigateur. Le profil de l'utilisateur est alors actualisé.

Les modes d'actualisation du nombre maximal de cadres et des attributs SHOW de chaque cadre restent à faire. Ils seront déterminés par l'éditeur de préférences de visualisation.

4.2.1.2.3. Méthodologie de modélisation

Hormis les noms des sous-domaines modélisés, la structure de ces profils est indépendante du projet et peut donc être réutilisée telle quelle. Seules les préférences par défaut changent. En principe, ces préférences ne devraient pas être très importantes compte tenu de la « personnalisabilité » de l'outil. Cependant, il est peu probable que des utilisateurs déroutés par le mode d'accès à l'information ou l'interface qui leur sont proposés s'intéressent aux

options de personnalisation de l'outil. C'est pourquoi nous considérons qu'il est important de leur proposer dès la première utilisation les options les plus adéquates possibles.

Nous avons vu plus haut que les préférences par défaut ont été définies en fonction des réponses des experts au cours des interviews. Il convient donc, lors de ces interviews, d'interroger les experts sur leurs démarches habituelles d'accès à l'information en plus de leur activité. L'objectif est de déterminer à la fois leur stratégie habituelle d'accès à l'information et les types d'informations les plus fréquemment recherchés. Il s'agit ici de définir des orientations générales moins précises que celles qui sont nécessaires pour la définition du modèle du SI.

Par exemple, dans le projet-pilote, nous avons clairement pu distinguer des sous-domaines d'ingénierie et de recherche. L'analyse des comportements des experts en situation de recherche d'information dans chacun de ces sous-domaines nous a permis de faire les assomptions mentionnées ci-dessus. Nous espérons que l'évaluation du système confirmera leur exactitude. Dans cette optique, nous proposons des pistes de caractérisation des domaines dont les démarches d'accès à l'information s'apparentent à celles de l'ingénierie ou de la recherche.

Les domaines d'ingénierie sont des disciplines opérationnelles dans lesquelles les individus ont des contraintes temporelles fortes. L'information recherchée doit être immédiatement utilisable. Pour cela, elle doit être aussi proche que possible de l'information à produire. Par exemple, un ingénieur qui réalise une étude de criticité pour la conception d'une installation va commencer par rechercher toutes les notes d'études de conception d'installations similaires. Ensuite, à chaque étape de son étude, il procède par analogie avec des documents portant sur des configurations similaires pour la même étape. Les autres types d'informations auxquelles il se réfère sont aussi bien définies. Il s'agit des standards et autres informations réglementaires. Il n'est pas dans ses habitudes d'en extrapoler la validité pour des situations de travail différentes⁹⁴.

Au contraire, les domaines de recherche s'intéressent à des types d'informations plus variables. Ces informations peuvent par exemple émaner de domaines connexes ou être des publications récentes. Les besoins en informations des chercheurs sont donc plus difficilement prédictibles que ceux des ingénieurs. Cependant, une définition « acceptable » et acceptée par

⁹⁴ Du moins, c'est le cas dans le domaine de la sûreté-criticité nucléaire.

les experts du SEC est la suivante : les chercheurs ont en général des contraintes de temps moins fortes et s'intéressent à l'ensemble de l'information disponible sur un sujet.

Par conséquent, une méthode de mise au point des préférences par défaut des utilisateurs est l'identification du domaine modélisé à l'une ou l'autre de ces définitions. Cette identification peut être faite a posteriori par le chargé de projet qui a interviewé l'expert. La dimension opérationnelle et réglementée ou au contraire vaste et parfois même mal définie de la discipline est évoquée pendant les interviews.

4.2.1.3. Le modèle conceptuel du Système d'Information

Le système d'information existant dans l'organisation avant le démarrage du projet est constitué d'un ensemble d'unités documentaires a priori non structurées et disparates. Nous cherchons ici à en déterminer un modèle permettant l'accès et le filtrage de ce système d'information. Nous donnons dans ce paragraphe des descriptions conceptuelle puis technique de ce modèle et décrivons ensuite une méthodologie permettant de le mettre au point.

4.2.1.3.1. *Description du modèle du SI*

Nous avons vu que la recherche d'information peut avoir lieu en fonction des éléments des diagrammes (il s'agit alors d'une requête sur le contenu de l'information recherchée). Elle peut également prendre la forme d'une recherche par champs (auteur, date, type de document). Le filtrage permet de sélectionner des types de documents ou des ensembles de documents se rapportant à un descripteur spécifique au domaine (voir stratégie d'accès à l'information cumulative).

Ce modèle doit permettre la description des unités informationnelles sous les points de vue des différents types d'utilisateurs et d'activités autour du domaine considéré. Pour cela, il s'appuie sur différents types de critères, et la description d'un élément dans ce modèle n'est pas unique.

L'identification des critères pertinents pour cette description est faite durant la phase de modélisation des connaissances et de définition des profils des utilisateurs. Ces critères sont donc fortement dépendants du domaine, voire des préférences des experts interviewés. Le modèle obtenu est donc différent dans chaque projet. Cependant, il possède une structure générique réutilisable qui est « remplie » à chaque nouveau projet. La méthode d'obtention de ces critères est précisée au paragraphe suivant.

Deux grandes catégories de critères ont été choisies pour ce modèle. Les critères de forme (ou critères de classification documentaire) permettent de décrire et de retrouver physiquement les documents, fichiers et autres « grains d'information » ; ce sont des champs classiques (titre, auteur, date, type de support...). Les critères de fond sont eux spécifiques au domaine (sous-domaine, niveau de difficulté, activités génériques utilisant cette information...) ; c'est grâce à eux que va s'organiser la recherche documentaire.

Ce modèle doit être compatible avec la modélisation des connaissances. Cela implique que les critères employés aient un sens vis à vis des éléments des diagrammes et que la correspondance entre ces derniers et le SI soit facilement opérationnalisable. Pour cela, nous avons choisi de présenter ce modèle sous une forme arborescente : le SI y est décrit sous la forme d'une série de critères qui peuvent prendre plusieurs valeurs, et chacune d'elles pointe vers les unités informationnelles correspondantes. Les éléments des diagrammes et les caractéristiques de l'utilisateur sont associés aux valeurs des différents critères. La recherche s'effectue alors sur l'intersection des ensembles d'unités correspondantes.

En pratique, dans le projet-pilote sur la maîtrise du risque de criticité, nous avons défini un modèle des concepts spécifique pour l'indexation conceptuelle du SI⁹⁵. Ce modèle regroupe un ensemble de concepts (et d'objets) contenus dans les diagrammes issus de la modélisation MKSM. Cependant, il est plus restreint et tient compte des critères de classification identifiés par les différentes actions de classification du fonds documentaire menées au SEC⁹⁶. Nous avons conservé ces classifications pour structurer l'information disponible. Nous avons ensuite fait correspondre à ce modèle l'ensemble des éléments du système d'information que nous avons répertoriés au SEC. La figure ci-dessous montre une partie du modèle du SI réalisé pour le projet-pilote.

⁹⁵ Nous ne faisons référence ici qu'à la partie « critères de fonds » du modèle du SI.

⁹⁶ Notamment, la principale notion qui entre dans le système de classification sémantique des éléments du SI au SEC est le type de matière fissile auquel ils se rapportent.

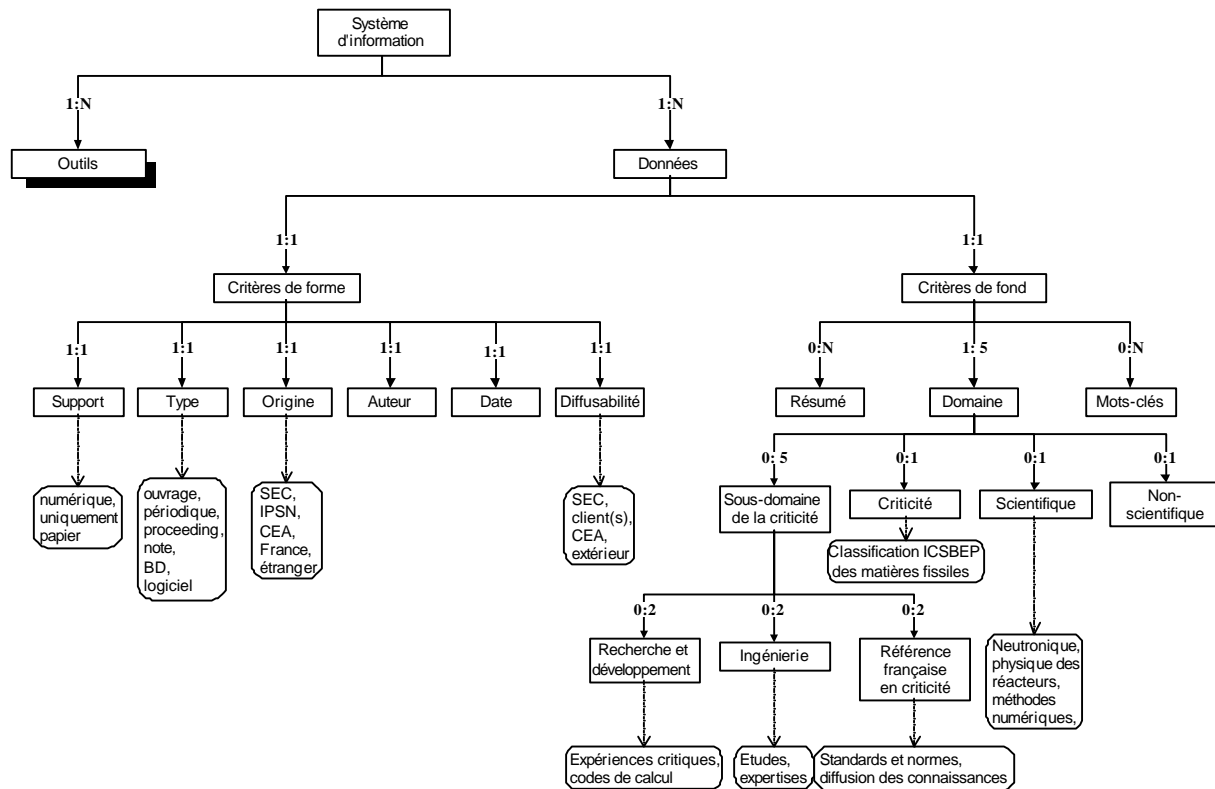


Figure 23. Les critères de classification des éléments de données du SI⁹⁷ du SEC.

Techniquement, le modèle est réalisé sous la forme d'une arborescence d'index des documents répondant à chacun des critères. Chacun d'eux est sélectionné par le LMCE. Dans le prototype réalisé pour le projet-pilote, les fonds documentaires sont au format pdf et sont donc indexés par des index spécifiques au moteur de recherche Acrobat Search. La sélection d'un tel index depuis une page HTML est notamment possible grâce à la technologie ActiveX, qui permet d'incorporer des objets actifs tels que des plug-ins Acrobat. Nous avons réalisé un tel plug-in capable de réaliser la sélection d'index et de lancer une requête et l'avons incorporé dans un contrôle ActiveX. Les différents objets de la fenêtre HTML peuvent interagir avec ce contrôle via le DOM Microsoft⁹⁸. Nous indiquons plus loin comment le LMCE (et plus précisément le système de règles) sélectionne les index correspondant au filtrage des informations.

⁹⁷ Le même type de classification a été réalisé pour les applications. Il est présenté dans le paragraphe traitant du projet-pilote. Nous n'avons pas jugé opportun de la faire figurer ici. En effet, si le sujet initial de la thèse concernait l'intérêt de la méthode MKSM dans le contexte de l'utilisation du système d'information, nous avons choisi de nous limiter à l'étude d'une démarche d'accès à l'information centrée sur l'utilisateur et la modélisation du contexte d'activité.

⁹⁸ C'est-à-dire le DOM tel qu'il est implémenté dans le Microsoft Internet Explorer®. Le DOM de Netscape Navigator® est sensiblement différent et n'est pas considéré ici.

4.2.1.3.2. *Méthodologie de modélisation*

L'indexation du fonds documentaire de l'organisation est réalisée en deux étapes. D'une part, il s'agit de définir tous les critères d'indexation pertinents pour la recherche d'information spécifique à l'activité modélisée. Une fois ces critères définis, il faut identifier les documents s'y rapportant.

4.2.1.3.2.1. **Identification des critères d'indexation**

Tous les critères identifiés dans la figure ci-dessus ne sont pas pertinents pour définir des index. En particulier, les auteur, date ou résumé d'un document ne peuvent être utilisés que pour une recherche par champs. Nous ne nous intéressons pas ici à ces critères. L'objectif de ce paragraphe est de généraliser la démarche de définition des critères d'indexation à d'autres projets en s'appuyant sur le modèle présenté pour le projet-pilote.

Concernant le support, seules les informations sur support numérique sont traitées ici. Nous n'avons pas cherché à définir le référencement de documents disponibles sur d'autres supports. Ce type de travail, qui s'apparente aux catalogues en ligne des bibliothèques nécessite la rédaction d'une classe d'éléments d'information supplémentaires (les références) visant à donner une description au format électronique des documents disponibles sous un autre format. Ce travail ne peut être fait par le chargé de projet en gestion des connaissances car il est extérieur au domaine⁹⁹. Toutefois, si ces références existent, elles sont traitées comme tous les autres éléments du système d'information.

L'identification des types d'informations les plus couramment utilisés dans le domaine peut être déduite des modèles de connaissances¹⁰⁰. Les modèles de type flots de données (i.e. le modèle d'activités MKSM) permettent de présenter les flux – et notamment les flux documentaires – qui circulent entre les activités. Les diagrammes correspondants identifient donc, dans des éléments de type flux, les types de documents qu'il est nécessaire d'indexer. Cette détermination des types de documents présente l'avantage de pouvoir être menée sans l'aide de l'expert.

Les autres modèles dynamiques, comme le modèle des tâches MKSM peuvent également faire référence à des types de documents. C'est le cas lorsque les diagrammes de tâches

⁹⁹ De plus, ce type de travail irait à l'encontre de la politique de l'unité où nous avons réalisé le projet-pilote, qui est de numériser et de rendre disponible électroniquement tous les documents papiers de son système d'information.

¹⁰⁰ Pour cela, il faut que la modélisation des connaissances ait été orientée en fonction de l'accès à l'information, comme cela est décrit au chapitre suivant.

mentionnent explicitement les types d'informations nécessités pour la réalisation de la tâche¹⁰¹. L'aide de l'expert est parfois requise pour déterminer plus précisément les types de documents utilisés pour chaque tâche. Il s'agit alors de lui poser la question « quels sont les types d'informations utilisés pour réaliser cette tâche ? » Cette méthode, plus coûteuse en temps pour l'expert et le chargé de projet, est plus précise que l'analyse des diagrammes d'activités, car plus proche de chaque tâche de l'expert.

D'autres types de documents peuvent exister et être utilisés dans le travail des experts sans qu'ils figurent explicitement sur les diagrammes. Ces documents peuvent être regroupés sous des appellations génériques. Pour cela, il est nécessaire d'obtenir une liste exhaustive des autres types de documents lors de la phase de classification documentaire décrite plus loin.

L'origine et la diffusabilité des documents sont des critères à définir avec les experts, qui connaissent les organisations externes avec lesquelles ils interagissent.

Concernant les critères de fonds, le sous-domaine d'activité est le premier critère permettant de dissocier l'utilisation des documents dans l'organisation. Il sert notamment au filtrage des éléments en fonction du sous-domaine dans lequel s'effectue la recherche d'information. Ce critère est obtenu grâce à la liste des sous-domaines modélisés.

La liste des domaines connexes est obtenue de la même manière que celle des types de documents.

Les mots-clés sont des descripteurs des éléments du système d'information spécifiques au domaine. Ces mots-clés correspondent à un champ sur lequel une recherche d'information peut s'effectuer. Cependant, nous avons noté que certains mots-clés sont plus fréquemment utilisés que d'autres. C'est le cas par exemple des différents types de matière fissile en criticité. Quasiment aucune recherche d'information n'a lieu sans la spécification de ce critère. C'est pourquoi, afin d'accélérer la recherche et de retrouver les documents pour lesquels ce champ n'est pas renseigné, nous avons jugé utile de considérer ce critère comme un critère d'indexation. Par définition, les principaux mots-clés du domaine sont ceux identifiés comme tels dans le modèle des concepts MKSM. Pour le projet-pilote, nous retenons la classification ICSBEP (International Criticality Safety Benchmark Evaluation Project) des matières fissiles pour déterminer ces critères.

¹⁰¹ La référence aux standards de criticité ou à différents types de notes techniques est par exemple fréquente dans le projet-pilote.

4.2.1.3.2.2. **Obtention des index**

Une fois la liste de ces index obtenue, il s'agit d'y faire correspondre les éléments du système d'information. Pour cela, le chargé de projet doit inventorier l'ensemble des éléments du système d'information disponibles électroniquement et déterminer, pour chacun d'eux, s'il est pertinent par rapport à chacun des critères. Il peut éventuellement s'aider de classifications existantes ou de l'avis des experts. Une fois cette classification obtenue, il réalise l'indexation de tous les documents correspondant à chaque critère de la liste dans un index spécifique.

Techniquement, le logiciel Acrobat Catalog destiné à indexer les fonds pdf¹⁰² permet la sélection de fichiers dans des répertoires différents. L'indexation multiple des documents sans recopie physique de ceux-ci est donc possible avec ce logiciel.

4.2.2. **Description fonctionnelle de l'accès à l'information avec le LMCE**

Les paragraphes précédents présentent les trois modèles conceptuels sur lesquels s'appuie le LMCE. Le but de cette partie est de mettre en lumière les différentes interactions entre les éléments de ces modèles dans le processus d'accès à l'information que propose cet outil. Il s'agit d'illustrer le fonctionnement de ce processus en en donnant une description fonctionnelle, sous forme de diagrammes d'activités (au sens de MKSM). Le processus décrit ici correspond principalement à un accès à l'information utilisant la stratégie spécifique (i.e. cumulative) proposée dans le paragraphe suivant.

Comme tout système d'accès à l'information, un livre de connaissances électronique propose une *fonction d'accès à l'information*. Nous employons volontairement le terme de fonction (plutôt que de fonctionnalité), par analogie avec la définition mathématique ou informatique de ce terme¹⁰³. Dans le cas du LMCE, l'argument est le problème d'information de l'utilisateur et la valeur de retour l'information supposée pertinente pour solutionner ce problème d'information. Le but d'un système d'accès à l'information est bien évidemment de réaliser les traitements nécessaires pour trouver la valeur de retour correspondant à l'information pertinente. Nous avons alors la représentation idéale :

LMCE

¹⁰² Le moteur livré avec le LCE doit, à terme être le moteur de recherche multilingue en texte intégral SPIRIT®, commercialisé par la société T-Gid. Pour l'instant, ce moteur ne peut cependant indexer et effectuer de recherche que dans des fichiers texte. Nous utilisons le moteur Acrobat Search® pour les fonds au format pdf.

¹⁰³ Une fonction est une entité qui transforme, moyennant certains *traitements*, un ou plusieurs *argument(s)* en une *valeur de retour*.



L'information pertinente est constituée d'un ensemble d'éléments du système d'information. Elle est indépendante du fonctionnement du système d'accès à l'information utilisé.

4.2.2.1. Le problème d'information comme argument de la fonction d'accès à l'information du LMCE

Nous avons vu au chapitre 2 qu'un problème d'information peut posséder plusieurs niveaux de définition et qu'une démarche d'accès à l'information s'appuie sur un besoin d'information, lequel est l'expression d'un problème d'information. Nous formulons ici l'hypothèse qu'un problème d'information est totalement décrit par ces deux caractéristiques indépendantes :

Problème d'information = niveau de perception + besoin d'information

Le niveau de perception du problème d'information de l'utilisateur est déterminé par son niveau d'expertise du domaine. Le fondement de cette hypothèse peut être vérifié ou infirmé par les tests de validation et le retour d'expérience sur l'utilisation du prototype. Toutefois, cela n'est pas fondamental dans le raisonnement ci-dessous. Le niveau de perception du problème d'information est une donnée que nous ne cherchons pas à modifier ici. Nous considérons qu'il est spécifié par la stratégie d'accès à l'information choisie par l'utilisateur ou par celle qui lui est assignée par défaut dans son profil. Dans le cadre de l'utilisation de la stratégie d'accès à l'information cumulative, le problème d'information est perçu soit au niveau exprimable soit au niveau formalisé¹⁰⁴.

Le besoin d'information est donné par le résultat de l'interaction avec le système. Quand l'utilisateur choisit d'utiliser la navigation, il est donné par la donnée de la tâche qu'il est en train d'accomplir et des éléments sélectionnés. Lorsqu'il choisit de poser directement une question, il est donné par la requête qu'il a formulée. Dans ce dernier cas, l'analyse ci-dessous est sans objet. Dans le premier cas, plusieurs composantes entrent dans la définition du besoin d'information :

la dimension professionnelle du profil de l'utilisateur, qui limite l'espace des informations pertinentes en fonction de ses niveaux d'expertise dans les différents sous-domaines,
la dimension informationnelle de son profil, qui permet d'affiner le filtrage de cet espace,

le modèle du domaine, qui contient les éléments de requête sélectionnés et identifie les tâches auxquelles ils correspondent.

Nous voyons ici que seuls les modèles de l'utilisateur et du domaine sont concernés par la définition du problème d'information. Ces deux modèles se situent donc en amont de la fonction d'accès à l'information du LMCE. Ils permettent de caractériser entièrement l'argument de cette fonction d'accès à l'information.

Le modèle de l'utilisateur définit le filtrage du système d'information spécifique au profil de l'utilisateur. Il identifie par conséquent l'espace des informations accessibles par l'utilisateur. Cet espace des possibles est constant au cours de l'utilisation du LMCE. C'est pourquoi il intervient en tant que ressource de l'utilisation du LCME.

Le modèle du domaine représente la partie variable de l'argument de la fonction d'accès à l'information du LMCE. Les éléments des diagrammes sélectionnés sont eux-même les résultats d'une autre fonction : la navigation hypermédia dans les diagrammes de connaissances. Cette navigation possède une *fonction pédagogique* (elle permet à l'utilisateur de comprendre l'activité représentée) et une *fonction documentaire* (elle lui présente des éléments à partir desquels il peut formuler son problème d'information). La fonction pédagogique de la navigation dans un hypermédia n'est pas discutée ici (voir par exemple [Tricot & al. 98] pour une bibliographie de cet aspect).

La fonction documentaire est ici traitée au niveau unitaire des éléments des diagrammes. Cette fonction s'appuie sur le formalisme du diagramme, sur le type de l'élément et sur son contenu (texte et valeur de reformulation). Elle fait correspondre une partie du système d'information à une composante particulière du besoin d'information exprimé par l'utilisateur.

Chaque formalisme diagrammatique possède un statut particulier, vis-à-vis du positionnement des modèles dans la méthode MKSM (voir chapitre 3). Les éléments de ces diagrammes possèdent en plus un *typage sémantique* spécifique. Certains de ces typages peuvent être associés aux contenus des éléments pour servir de base à la formulation du besoin d'information de l'utilisateur. Dans ce cas, les éléments des diagrammes correspondent à des critères d'indexation particuliers. Dans le cas contraire, ces éléments sont des composantes de

¹⁰⁴ Dans ce cas, nous considérons que l'utilisateur simule le niveau exprimable, soit pour gagner du temps sur la formalisation de sa requête, soit pour éviter de taper celle-ci au clavier.

la requête envoyée au moteur de recherche. Nous détaillons plus loin l'utilité de chacun des formalismes MKSM pour la présentation de ces critères.

4.2.2.2. Position des autres éléments de la modélisation

Le modèle de l'information et les préférences de visualisation de l'utilisateur ne sont pas des arguments de la fonction d'accès à l'information. Ces éléments permettent l'accomplissement cette fonction sans être consommés par celle-ci. Ils sont donc considérés comme des ressources (au sens du modèle d'activité MKSM).

4.2.2.3. La fonction d'accès à l'information

Le schéma suivant présente une vue globale de la fonction d'accès à l'information avec le LMCE. Il positionne les modèles conceptuels autour de cette activité.

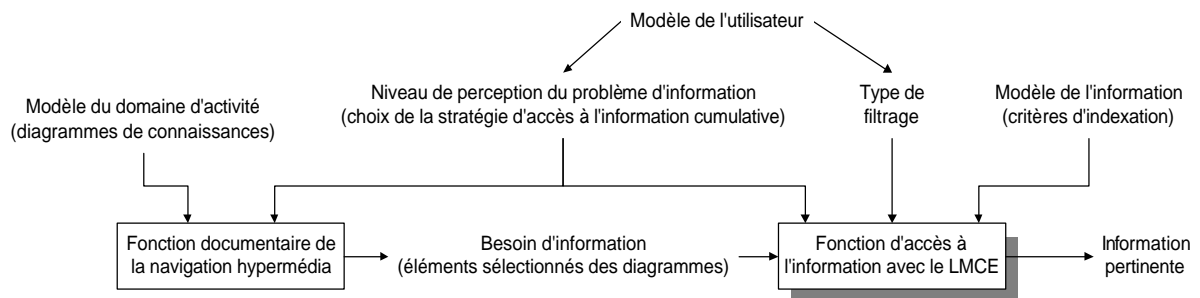


Figure 24. Position des modèles dans la fonction d'accès à l'information du LMCE. Les préférences de visualisation de l'utilisateur ne sont pas représentées. Elles interviennent en tant que ressources des deux activités.

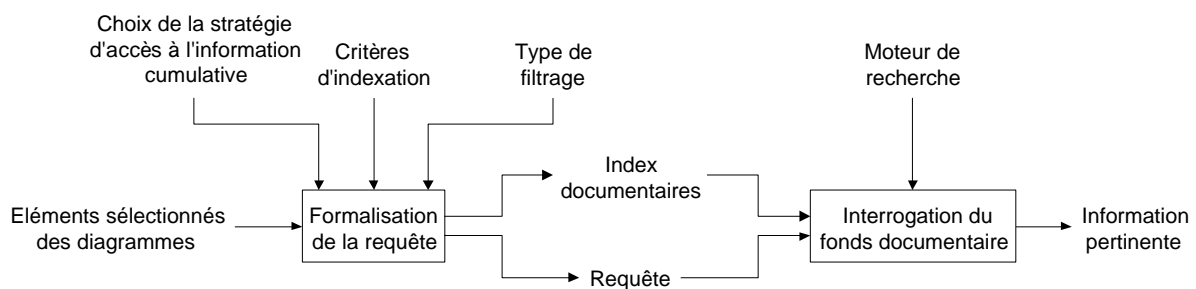


Figure 25. Détails de la fonction d'accès à l'information. L'interrogation consiste en la sélection des index et l'envoi de la requête à un moteur de recherche qui apparaît en tant que ressource de cette activité. L'activité de recherche d'information dans le corpus par ce moteur n'est pas représentée.

Sur ces diagrammes, il apparaît que les modèles de connaissances interviennent dans la fonction documentaire de la navigation hypermédia en permettant de sélectionner les éléments des diagrammes pertinents par rapport au problème d'information de l'utilisateur. Le modèle

de l'utilisateur intervient également dans cette activité, ainsi que dans celle de formalisation de la requête documentaire. Le modèle de l'information n'intervient que dans cette dernière activité.

4.2.3. Les outils d'expression du besoin d'information

Les paragraphes ci-dessus décrivent les modèles conceptuels et la démarche d'accès à l'information sur lesquels s'appuie le LMCE. Ces modèles conceptuels regroupent les données nécessaires au fonctionnement de l'outil d'accès à l'information. Nous présentons dans cette partie les deux principaux éléments qui favorisent cet accès à l'information. Le premier est un outil conceptuel. Il s'agit de la stratégie spécifique d'accès à l'information, qui propose une aide à l'expression du besoin d'information. Le second est le moteur de règles. Cet outil fonctionnel permet, à partir de l'expression d'une requête, un accès à l'information tirant parti de l'indexation conceptuelle du système d'information de l'organisation modélisée.

4.2.3.1. La stratégie d'accès à l'information cumulative

L'objectif est d'expérimenter une utilisation combinée des stratégies de parcours des diagrammes de connaissances et d'interrogation. La navigation dans les diagrammes de connaissances permet à l'utilisateur de passer en revue de nombreux éléments qui peuvent être pertinents pour la formulation de son besoin d'information. Il s'agit de lui permettre de sélectionner dans plusieurs diagrammes les éléments qu'il juge utiles et de spécifier des relations entre ces éléments pour former une requête documentaire complexe.

Cette stratégie d'accès à l'information spécifique au LMCE est baptisée *cumulative*. Dans ce paragraphe, nous justifions le choix d'une telle stratégie, analysons les tâches de l'utilisateur accédant à l'information par cette stratégie et définissons les éléments d'interface nécessaires pour la mettre en place.

La requête ainsi définie permet à l'utilisateur d'obtenir une formalisation de son besoin d'information, au même titre que ce que permet une interrogation directe. Cependant, elle peut s'appliquer à différents niveaux de perception du problème d'information, ce que ne permet pas l'interrogation directe¹⁰⁵. L'intérêt de cette stratégie d'accès à l'information, par rapport à une stratégie analytique, est double.

¹⁰⁵ Nous avons identifié au chapitre 2 deux caractéristiques du problème d'information : son niveau de perception, qui peut être ressenti, conscient, exprimable ou formalisé, et le besoin d'information (i.e. la description formelle de l'information capable de combler le « trou dans le système de connaissances de l'individu »).

Elle permet d'aider l'utilisateur dans l'expression de son besoin. Lorsque le problème d'information est perçu au niveau conscient, l'utilisateur est capable de le cerner et d'identifier les concepts pertinents pour sa recherche. En revanche, il n'est pas capable d'exprimer sa requête, même en langage naturel. La navigation dans la modélisation multi points de vue lui évite cette tâche en mettant directement à sa disposition l'ensemble des concepts utilisés dans le domaine. Il identifie alors ceux qui lui semblent décrire son besoin d'information et accède ainsi directement à une expression de ce besoin sous une forme proche du langage naturel, donc aisément compréhensible. Son problème d'information passe alors du niveau conscient au niveau formalisé.

Une fois le besoin exprimé, la stratégie cumulative d'accès à l'information aide l'utilisateur à formaliser ce besoin. Selon Gary Marchionini, la différence entre les niveaux exprimable et formalisé réside dans le passage de l'expression de ce besoin en langage naturel à un langage d'interrogation formel. Pour les utilisateurs peu familiarisés avec les langages d'interrogation, cette stratégie réalise cette traduction de l'expression du besoin d'information sans que l'utilisateur ait à faire l'effort de passage de l'un à l'autre de ces niveaux de perception du problème d'information.

Les points ci-dessus nous fournissent une définition des niveaux d'expertise du domaine en fonction des niveaux de perception du problème d'information : pour le LMCE, un utilisateur apprenant perçoit son problème d'information au niveau conscient. Aux utilisateurs novices correspond alors le niveau ressenti et aux experts celui d'exprimable ou de formalisé. Ces deux derniers niveaux sont distingués par l'expertise en outils d'accès à l'information du profil de l'utilisateur.

Un utilisateur novice est donc incapable d'identifier les concepts pertinents pour son problème d'information. C'est pourquoi la fonctionnalité d'accès à l'information ne lui est pas proposée, tandis qu'elle est proposée aux utilisateurs ayant des niveaux de perception supérieurs. La stratégie cumulative est proposée aux utilisateurs apprenants. Quelle que soit leur expertise des outils d'accès à l'information, elle leur permet au moins d'exprimer leur besoin d'information.

Elle est également utile aux experts des sous-domaines de recherche non familiarisés avec les langages d'interrogation. En effet, nous avons précédemment défini leurs besoins d'informations comme exhaustifs et centrés sur un sujet particulier. La recherche d'information ne peut donc pas se faire par champs. Elle s'appuie sur une requête complexe

qui doit être formulée dans le langage d'interrogation du moteur de recherche utilisé. Cette formalisation est réalisée par le système de règles.

4.2.3.2. Le système de règles

Une fois les éléments de définition du problème d'information de l'utilisateur identifiés et leur organisation spécifiée par l'utilisateur, il s'agit de formaliser la requête interrogative correspondant à ce problème. Dans les cas les plus simples¹⁰⁶, ce travail de formalisation de la requête est réduit à la juxtaposition des composantes du problème d'information. Dans des cas plus complexes, deux difficultés sont identifiées.

La première est de conserver l'articulation des éléments de la requête tout en respectant le langage d'interrogation utilisé par le moteur de recherche. Dans un premier temps, le respect de la syntaxe du langage peut poser problème pour les langages évolués, les plus contraignants d'entre eux étant les langages naturels. Cette difficulté est peu importante pour la conception du LMCE, car le langage utilisé (VDK) n'est pas d'une grande complexité. Toutefois, dans la perspective d'une interrogation du fonds par SPIRIT®, cette tâche deviendrait nettement plus complexe. Elle n'est pas traitée ici. Dans un second temps, les syntaxes des différents langages de requêtes énumérés au chapitre deux ne coïncident pas nécessairement avec l'articulation des éléments spécifiée par l'utilisateur. Il s'agit alors de traduire les relations spécifiées dans l'expression du besoin par les opérateurs du langage de requêtes correspondants, en réorganisant au besoin les éléments de la question.

Une fois le besoin d'information formalisé en fonction du langage d'interrogation, la seconde difficulté à résoudre est la sélection des index pertinents pour la recherche dans le cas d'une interrogation multi-bases. Il est important de noter que la recherche multi-bases se conçoit ici à la fois pour l'indexation différenciée d'un même corpus documentaire en fonction des critères définis plus haut et dans sa véritable acception. Le fait de pouvoir interroger plusieurs index, éventuellement à l'aide de plusieurs moteurs de recherche permet de consulter l'ensemble du système d'information de l'organisation à partir du LMCE¹⁰⁷.

¹⁰⁶ Moteur d'indexation et de recherche utilisant un langage de requêtes booléen et une indexation simple de la base documentaire, c'est-à-dire que toute la base est accessible à partir d'un seul index.

¹⁰⁷ Dans cette étude, nous nous limitons à l'accès à un fonds documentaire au format pdf. Il ne s'agit à notre sens pas d'une restriction mais d'un exemple d'application des spécifications du LCE à un format effectivement rencontré dans le projet-pilote. Le fait que ce format soit propriétaire et donc difficilement modifiable nous conduit à supposer que ce qui est possible avec un fonds pdf l'est également avec un fonds disponible sous autre format.

Nous avons vu aux paragraphes précédents que dans le cas du LMCE, plusieurs éléments sont à prendre en compte. Le filtrage s'effectue par exemple en fonction de certaines caractéristiques du profil de l'utilisateur, tandis qu'une analyse des termes de la question est également nécessaire pour identifier les critères d'indexation. Dans ce cas, ces termes doivent être supprimés de la question et les index correspondants sélectionnés. Les termes restants sont alors réorganisés en fonction de chacun des termes supprimés. Le système peut alors avoir à adresser des requêtes différentes aux différents index.

Pour illustrer cet envoi de différentes requêtes, nous considérons une requête initiale booléenne de type $(t_1 \text{ ET } c_1) \text{ OU } (t_2 \text{ ET } c_2)$, où t_1 et t_2 sont des termes non significatifs pour l'indexation et c_1 et c_2 sont des critères d'indexation. Nous appelons X l'ensemble des documents contenant le terme x , indépendamment du fait que x est ou non un critère d'indexation. Cette requête revient à rechercher les éléments de $(T_1 \cap C_1) \cup (T_2 \cap C_2)$, soit la réunion disjointe de $T_1 \cap T_2 \cap (C_1 \cap C_2)$, $T_1 \cap C_1$ et $T_2 \cap C_2$.

Nous justifions l'exclusion mutuelle de ces trois sous-ensembles par le fait que le LMCE n'a aucun contrôle sur les résultats de la requête. Il doit donc réaliser le tri des documents par ordre de pertinence a priori. D'où l'intérêt de positionner $T_1 \cap T_2 \cap (C_1 \cap C_2)$ en premier élément recherché. Pour éviter les réponses redondantes, il faut ensuite exclure les documents appartenant aux ensembles déjà explorés.

La recherche avec exclusion mutuelle s'écrit alors : $T_1 \cap T_2 \cap (C_1 \cap C_2)$; $T_1 \cap C_1 / (T_1 \cap T_2 \cap (C_1 \cap C_2))$; $T_2 \cap C_2 / (T_1 \cap T_2 \cap (C_1 \cap C_2))$. On peut encore écrire le deuxième ensemble (le troisième étant son symétrique) : $(T_1 / T_2) \cap (C_1 / C_2)$.

Trois requêtes sont donc nécessaires : il s'agit de retrouver, dans les documents respectivement indexés par $(c_1$ et $c_2)$, c_1 seul et c_2 seul, les termes $(t_1$ et $t_2)$, $(t_1$ et non $t_2)$, (non t_1 et $t_2)$.

L'intérêt d'un moteur de règles pour la résolution de ce type de problème est évident. Ce type d'outil permet les inférences logiques du type de celles présentées ci-dessus. En outre, il permet la spécification de conditions complexes comme celles qui président à la formalisation d'une requête avec le LMCE.

Les choix techniques effectués ici privilégient la simplicité de réalisation. En effet, le langage SDK utilisé par le moteur Acrobat Search® est quasiment un langage booléen, et l'articulation des composantes du besoin d'information dans le formulaire dédié à

l'interrogation cumulative suit à peu près ce langage. En conséquence, les règles à appliquer sont de simples comparaisons et les traitements sont des concaténations de chaînes de caractères. Ces traitements sont aisément implémentables dans des scripts du prototype et sont donc réalisés ici de cette manière. Le moteur de règles est donc un script écrit en JSCRIPT inclus dans le formulaire HTML d'interrogation cumulative. Il permet donc un post-traitement en deux étapes du contenu de ce formulaire. La première étape permet la visualisation de la requête « théorique » (i.e. ne tenant pas compte des critères d'indexation). L'utilisateur peut alors visualiser cette requête et la modifier s'il le désire. La séparation des termes et des index a lieu lors de l'envoi de la requête.

Dans la perspective de l'utilisation d'un moteur de recherche plus complexe, de la personnalisation du formulaire d'interrogation ou de la mise au point d'un atelier de génération de ce type de systèmes, un outil permettant la programmation à base de règles¹⁰⁸ serait sans doute nécessaire. Un tel outil améliorerait les performances, la compréhensibilité et la réutilisabilité de la fonctionnalité de formalisation de la requête du prototype. Les améliorations envisageables par l'incorporation d'un tel outil dans le LMCE sont discutées dans le paragraphe concernant les évolutions possibles de l'outil du chapitre suivant. La prospection d'un tel outil parmi l'offre disponible n'est pas traitée ici.

4.2.4. Identification des tâches et mise au point des composants de l'interface

À partir des modèles et des outils définis aux paragraphes précédents, nous pouvons identifier les différentes tâches à réaliser et proposer des éléments d'interface spécifiquement destinés à ces tâches. L'objectif de cette phase de la conception est une démarche de définition ergonomique des éléments de l'interface. Comme nous l'avons indiqué au chapitre trois, nous nous référons aux critères de conception ergonomiques énoncés dans [Bastien & al. 98] pour cette phase.

Deux fonctions générales sont identifiées pour le LMCE : la formation et l'accès à l'information. Dans les paragraphes suivants, nous décrivons les tâches et les éléments de l'interface qui permettent de réaliser ces fonctions. Toutefois, d'autres fonctions ponctuelles (déclaration du profil, connexion) sont également à réaliser. Les tâches et les éléments d'interface correspondants sont décrits dans une troisième partie. Dans une quatrième partie,

¹⁰⁸ Par exemple, un moteur de règles en Prolog.

nous passons en revue l'ensemble des critères ergonomiques pour chaque composant de l'interface.

4.2.4.1. L'utilisation du LMCE pour la formation

L'étude des processus d'apprentissage via l'utilisation d'outils tels que l'hypermédia est un domaine complexe qui s'« *efforce d'associer études fondamentales et applications* » ([Rouet 98]). Nous n'entendons pas développer d'étude dans le domaine de l'apprentissage. C'est pourquoi nous n'avons pas réalisé de modélisation des tâches pour cet aspect. Nous nous appuyons sur des travaux réalisés par des auteurs compétents dans ce domaine. Notamment, [Tricot & al. 98] présentent une synthèse des méthodes d'aide à l'apprentissage en environnement hypermédia, classées en fonction des types d'apprentissages pour lesquels elles sont pertinentes.

Le retour d'expérience sur les projets de gestion des connaissances prouve que le livre de connaissances est un bon outil de présentation du domaine. L'objectif est de l'utiliser pour permettre la découverte et la compréhension globale de toutes les relations entre les différents concepts du domaine. L'identification des éléments d'interface du LMCE pertinents pour les utilisateurs novices s'appuie sur l'emploi d'une stratégie de parcours pour l'exploration du livre de connaissances. Le paradigme est alors la présentation au format hypermédia de ce livre de connaissances (i.e. le LCH) et l'activité de l'utilisateur est la navigation hypermédia dans ce LCH.

Nous nous inscrivons ici dans le contexte d'un « apprentissage par exploration » de [Tricot & al. 98]. Ces auteurs identifient trois types d'aides pertinents pour ce type d'apprentissage.

4.2.4.1.1. *Les aides à la gestion de la tâche*

Pour ces auteurs, « *il s'agit essentiellement d'aider le sujet à se représenter de façon opérationnelle le but qu'il poursuit.* » Dans le cas d'une utilisation du LCH par un utilisateur novice, nous avons vu que celui-ci se situe dans un contexte d'apprentissage par exploration. Il nous semble alors que l'exploration de l'hyperdocument constitué par le LCH est le but en soi de l'utilisation de cet outil. Nous n'avons donc pas tenu compte de cette recommandation pour ce type d'utilisation.

4.2.4.1.2. Les aides à la sélection d'informations

Deux éléments sont ici à prendre en considération. Il s'agit de permettre à l'utilisateur d'identifier facilement les éléments actifs (i.e. les ancres des liens hypermédia) et de « *définir des menus d'accès simples et peu profonds* » pour la sélection de ces liens. Ces éléments sont destinés à faciliter les deux premières tâches du modèle EST (évaluation, sélection, traitement) de [Rouet & Tricot 95]¹⁰⁹.

Concernant la particularisation des éléments des diagrammes à la source des liens hypermédiés, il convient de distinguer les différentes catégories de liens inter-diagrammes que nous avons identifiées dans le projet de DEA antérieur à cette thèse ([Médini 97]). Les liens inter-diagrammes peuvent être intra-modèles (et dans ce cas intra ou inter formalismes) ou inter-modèles. Les liens intra-diagrammes ne sont pas traités ici puisqu'ils relient des éléments du même ~~nom~~ de l'hyperdocument.

Pour les liens intra-modèles, la particularisation des ancres des liens est prévue par le formalisme pour les liens intra-formalismes ou par la méthode MKSM pour les liens inter-formalismes. Par exemple, une activité, une tâche ou un concept, lorsqu'ils sont détaillés dans un autre diagramme du même formalisme, apparaissent en ombré sur les diagrammes sources. Les liens inter-formalismes comme ceux existant entre scénario et processus du modèle du domaine MKSM font explicitement référence au nom du diagramme cible dans le diagramme source.

De la même façon, il appartient au rédacteur du livre de connaissances de particulariser les ancres des liens inter-modèles dans les diagrammes sources. Dans le projet-pilote, nous avons choisi de particulariser les éléments des diagrammes à la source de liens inter-modèles par des annotations, comme cela est présenté dans le diagramme ci-dessous.

¹⁰⁹ La troisième tâche, qui est le traitement de l'information présentée, est facilitée par la clarté de présentation de l'information mais n'est pas spécifique aux environnements hypermédiés. Dans notre cas, l'information présentée est le diagramme MKSM, dont la clarté est avérée par le retour d'expérience des projets.

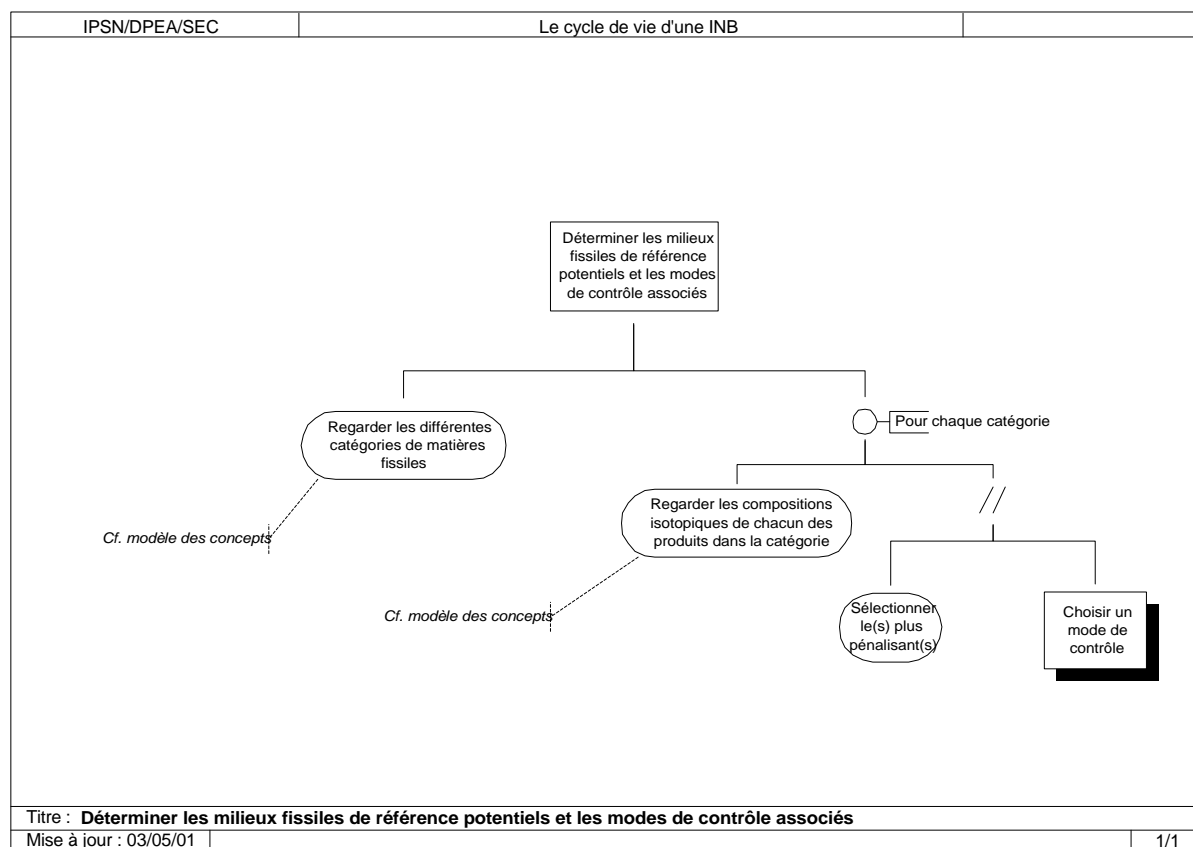


Figure 26. Exemple de diagramme de tâches comportant des liens intra et inter modèles. Les annotations signifient que des diagrammes de concepts font référence au contenu des éléments correspondants, et l'ombre derrière la tâche « Choisir un mode de contrôle » signifie qu'elle est décomposée dans un autre diagramme de tâches.

Une fois un élément d'un diagramme identifié à la fois comme pertinent par rapport au but de sa navigation hypermédia et comme une ancre d'un lien, l'utilisateur va sélectionner ou non un lien à partir de cet élément. Pour l'aider à faire son choix, il faut lui fournir une information appropriée sur le ou les liens atteignables depuis cet élément.

Pour [Bastien & al. 98], « *le texte du lien devrait être redondant avec le titre de la page qu'il fait apparaître.* » Cette redondance, est réalisée dans le LCH pour toutes les relations intra-modèles. Il arrive cependant que pour certaines relations inter-modèles, elle ne soit pas respectée. C'est notamment le cas pour les liens dont la cible devrait être un sous-élément d'un diagramme et non le diagramme entier¹¹⁰. Cette contrainte est due au fait que les diagrammes sont exportés sous forme d'images, et que l'accès à un élément isolé de leur structure est impossible. De plus, le paradigme systémique sur lequel se fonde MKSM veut que chaque élément du diagramme se définisse en fonction des autres. On peut donc

¹¹⁰ Par exemple, lorsqu'une activité fait référence à un type de matière fissile, celui-ci apparaît dans un diagramme comme sous-concept du concept « matières fissiles ».

s'interroger sur l'intérêt de visualiser d'un élément isolé, même s'il possède lui-même des sous-éléments. Une solution intéressante à ce problème serait sans doute la recomposition des graphes en fonction de l'élément sur lequel il s'agit de focaliser. Nous n'avons pas exploré cette voie dans ce travail car elle nécessite un calcul dynamique des arcs des graphes, ce qui nécessite un travail supplémentaire considérable, en plus d'un nouvel outil de production des diagrammes.

Par ailleurs, les relations entre les diagrammes MKSM sont sémantiquement typées. Les types des relations dépendent des formalismes des diagrammes sources et cibles (voir [Médini 97] pour la liste de ces types de relations). La mise au format hypermédia de ces liens ne doit pas impliquer la perte de ces informations sémantiques. C'est pourquoi nous avons choisi de faire figurer la sémantique du lien et le type du diagramme cible à la sélection du lien. Ainsi, l'utilisateur connaît à l'avance non seulement la destination, mais également l'inférence à réaliser en sélectionnant un lien entre deux diagrammes.

En outre, dans la méthode MKSM, plusieurs relations vers des diagrammes de types différents peuvent être à la source d'un même nœud. Il faut donc insérer entre la phase de sélection de l'élément du diagramme et l'activation d'un des liens à partir de cet élément, une phase supplémentaire de sélection du lien à activer. Cela est réalisé par le recours à un élément d'interface que nous avons baptisé « multiplexeur de liens », et qui permet l'affichage des informations relatives à chacun des liens partant de l'élément sélectionné.

Techniquement, il s'agit de proposer à l'utilisateur plusieurs liens à partir de la même ancre, ce que ne permet pas le langage HTML. Comme dans le prototype de Jean-Marc Charlot, ce multiplexeur utilise un script pour récupérer les données concernant les liens (éléments <LIEN> du nœud correspondant) dans le fichier XML du diagramme affiché. Cela est classiquement réalisé en HTML par un élément de type « DIV », qui permet de définir un rectangle d'affichage spécifique. Ce rectangle ne devient visible qu'à la réception de l'évènement « onclick » sur l'élément du diagramme correspondant. À la différence avec le prototype de Jean-Marc charlot, ces multiplexeurs sont définis dans la feuille de style XSL associée aux diagrammes et non plus « en dur » dans le fichier HTML. L'incorporation des données sur les liens est faite par concaténation dans le script. Ce multiplexeur n'a pas encore été réalisé pour ce prototype. Un exemple de syntaxe de code HTML réalisant ce type d'opération est toutefois disponible dans [Pardi 99], pp. 108-109. La figure ci-dessous montre un tel élément tiré du prototype de Jean-Marc Charlot.

Dès lors, une nouvelle itération d'évaluation du contenu de l'information sur les liens proposés, de sélection du lien approprié et de traitement (click sur le lien) a lieu. D'après ce qui précède, pour chacun de ces liens, il faut donc à nouveau indiquer dans ce multiplexeur le type et le titre de chacun des diagrammes cibles. Par souci d'homogénéité, ce mécanisme de parcours est aussi valable à partir des éléments des diagrammes ne comportant qu'un seul lien inter-diagrammes. Nous montrons plus loin que cette boîte est également utilisée pour l'accès à l'information.



Figure 27. Élément de sélection des liens à partir d'une activité décomposée dans le prototype de Jean-Marc Charlot. Cet élément est également utilisé pour l'accès à l'information.

4.2.4.1.3. *Les aides à la compréhension*

Dans cette section, [Tricot & al. 98] préconisent notamment de favoriser la compréhension des liens. Cela est fait avec l'explicitation de la sémantique des liens décrite ci-dessus. En complément de l'explicitation de ces liens, il s'agit ici de favoriser la compréhension des nœuds de l'hypertexte. Pour cela, deux types d'outils sont disponibles dans le LMCE : les cartes conceptuelles et le bandeau de titre.

4.2.4.1.3.1. **Les cartes conceptuelles**

D'un point de vue plus général, comme nous l'avons déjà souligné, nous souhaitons favoriser la compréhension de l'information par sa remise en contexte. Cela est vrai pour l'information relative au domaine étudié, que nous recontextualisons grâce aux diagrammes de connaissances. Cela est également vrai pour l'information contenue dans ces diagrammes pour laquelle la méthode MKSM propose des outils spécifiques.

L'objectif est de permettre à l'utilisateur de se situer dans le domaine et de se diriger vers les points qu'il souhaite approfondir. Nous avons vu au chapitre 2 que la présentation de cartes conceptuelles se prête à cet objectif.

La méthode MKSM fournit un éventail de cartes de ce type (arbre fonctionnel, dictionnaires des acteurs, des concepts et des flux) adaptées aux formalismes diagrammatiques utilisés. Ce sont donc des éléments que l'interface doit proposer aux utilisateurs novices, en même temps que les diagrammes de connaissances dans lesquels ils naviguent.

Le fait de disposer de plusieurs types de cartes conceptuelles pertinentes en fonction du formalisme diagrammatique visualisé suppose deux choix dans la présentation de ces cartes. Il s'agit d'une part de déterminer le nombre de cartes qu'il est souhaitable d'afficher à l'écran, et d'autre part d'identifier le mode de succession de ces cartes à l'écran le plus pertinent. À titre d'exemple, dans le prototype réalisé par Jean-Marc Charlot, une seule carte conceptuelle est présentée à la fois à l'utilisateur. L'alternance de ces cartes conceptuelles est synchronisée avec le changement de type de diagramme visualisé. Elle est aussi possible par un choix explicite de l'utilisateur.

Nous n'avons trouvé aucun travaux dans la littérature concernant la détermination du nombre de cartes à présenter à l'utilisateur. Nous formulons donc l'hypothèse, qui reste toutefois à valider, que la présentation de plusieurs cartes conceptuelles pour l'accès à un même espace d'information selon des points de vue différents peut perturber un utilisateur. À plus forte raison, pour un utilisateur novice, elle accroît la quantité d'information à assimiler et participe d'une éventuelle « surcharge cognitive ». À ce titre, nous nous limitons dans le LMCE à un seul type de carte conceptuelle pour l'accès aux diagrammes de connaissances.

Concernant la synchronisation de la carte et des diagrammes, nous voyons que ce choix n'est pas trivial. Il pose notamment la question du statut de la carte conceptuelle comme élément de présentation dynamique de l'espace parcouru. À quoi sert en effet une carte si elle varie à la vitesse à laquelle le déplacement a lieu ? Peut-elle tout de même servir de « référentiel mouvant » pour se repérer ? Hormis la supposition que l'utilisateur peut, après un temps d'adaptation, comprendre un tel enchaînement de cartes, nous n'avons pas d'élément objectif de réponse à ces questions. En revanche, nous pouvons objectivement affirmer qu'il ne sert à rien de visualiser une carte d'un espace où l'utilisateur ne se trouve pas. Il nous semble donc plus opportun de considérer comme inenvisageable la perte de cohérence entre la carte et le diagramme. La synchronisation de ces éléments est donc assurée par le système.

Techniquement, les cartes conceptuelles sont des documents XML qui peuvent être générés automatiquement à partir de la collection de fichiers XML correspondant aux diagrammes présentés plus haut. Ces cartes sont affichées, via une feuille de style XSL spécifique, dans un cadre HTML séparé du diagramme. La synchronisation entre ces deux éléments de l'interface

se fait en HTML dynamique dans des scripts utilisant le modèle objet des documents (DOM) Microsoft®. Cette technique est décrite plus en détails et illustrée sur un élément d'interface déjà réalisé plus loin dans ce chapitre.

4.2.4.1.3.2. Le bandeau de titre

Pour ne pas être perdu dans l'hypermédia, l'utilisateur doit savoir non seulement où il va, mais également où il se trouve. Pour cela, il doit avoir connaissance du titre du diagramme qu'il est en train de visualiser. D'autres informations peuvent être utiles, comme le formalisme de ce diagramme ou une trace des diagrammes précédemment visualisés. Or, l'exportation en HTML des diagrammes Visio se fait après la suppression du « cartouche » présentant certaines de ces informations (i.e. le titre et le formalisme)¹¹¹. Le diagramme visualisé ne comporte donc plus ces informations. Il est alors nécessaire de faire apparaître ces informations à l'écran. Pour cela, nous avons recours à un « bandeau de titre ». Cet élément d'interface nous sert à d'afficher toutes les informations contextuelles indiquées plus haut. Nous verrons plus loin que cet élément est également présent à l'écran pour les autres types d'utilisation du LMCE. C'est pourquoi l'activité¹¹² générale de l'utilisateur (i.e. navigation, recherche d'information) est également indiquée.

Ce bandeau est réalisé sous la forme d'un cadre HTML. Comme il est destiné à être affiché en haut de l'écran (voir paragraphe « aspects globaux de l'interface »), il présente les informations nécessaires sur toute la largeur de l'écran. Il est affiché dans un cadre spécifique, qui permet d'éviter son rechargement à chaque étape de navigation. Il est cependant actualisé à chaque étape. Les aspects techniques d'actualisation des différents cadres de la page sont détaillés plus loin.

4.2.4.2. L'utilisation du LMCE pour l'accès à l'information

Le LMCE permet deux stratégies d'accès à l'information. L'utilisateur peut interroger directement le fonds documentaire ou construire une requête par navigation dans les diagrammes de connaissances. La seconde stratégie est spécifique à l'outil, et est décrite au paragraphe suivant. L'analyse des tâches et la définition des composants de l'interface décrites ensuite sont communes aux deux stratégies.

¹¹¹ La suppression du cartouche est destinée à diminuer les défilements nécessaires à la consultation d'une image. Le cartouche est destiné à entourer le dessin imprimé sur une feuille au format A4. Pour les « petits » dessins, cela permet de réduire notablement la taille de l'image.

¹¹² Nous préférons ici parler d'activités plutôt que de tâches car nous nous intéressons à une utilisation globale du système et non au détail des actions de l'utilisateur.

4.2.4.2.1. Description des tâches

Nous avons modélisé les tâches à réaliser pour l'accès à l'information avec le LMCE en utilisant le formalisme des tâches MKSM. À notre sens, ce formalisme permet de visualiser graphiquement les différents éléments de contrôle des tâches, comme les séquences et les alternatives. Nous mentionnons cependant les préconditions (au sens de MAD) suivantes :

un diagramme de connaissances est affiché. Il s'agit soit du diagramme principal, soit du dernier diagramme exploré. Il appartient au chargé de projet de déterminer le diagramme par défaut le plus pertinent pour chaque profil d'utilisation.

le niveau de perception du problème d'information est au minimum conscient.

Trois diagrammes de tâches sont nécessaires pour décrire l'accès à l'information avec le LMCE. Ils sont présentés ci dessous.

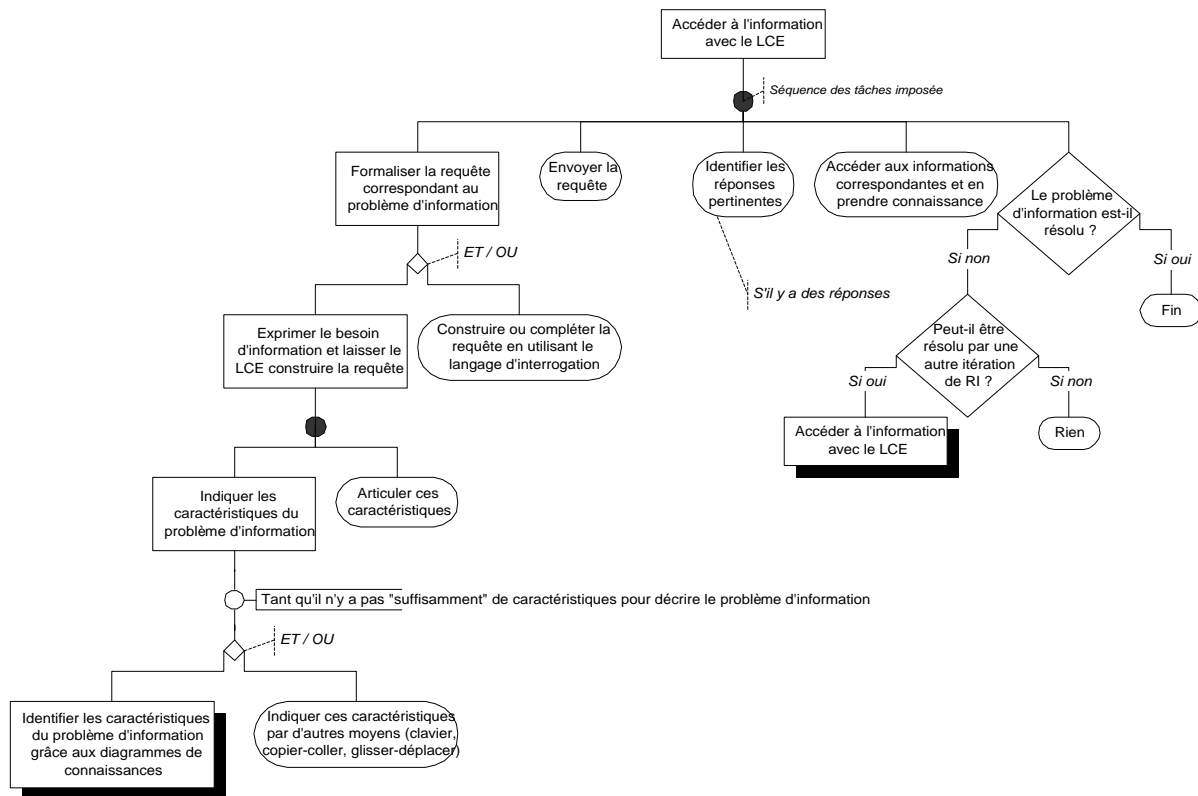


Figure 28. Les tâches correspondant aux différentes stratégies d'accès à l'information avec le LMCE.

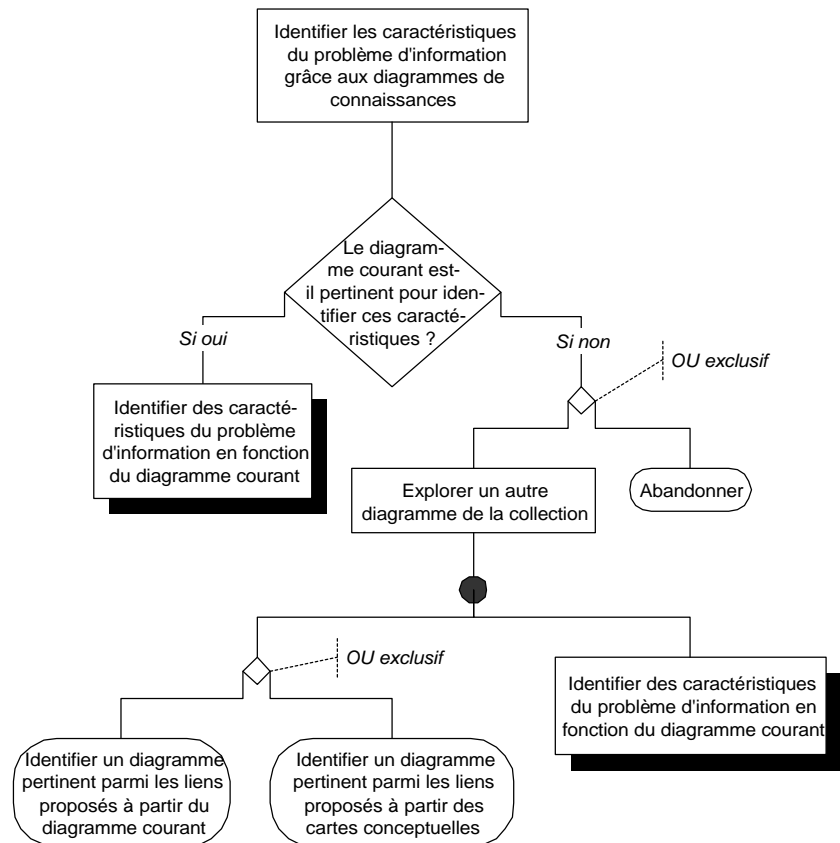


Figure 29. Détails de l'exploration des diagrammes centrée sur l'accès à l'information.

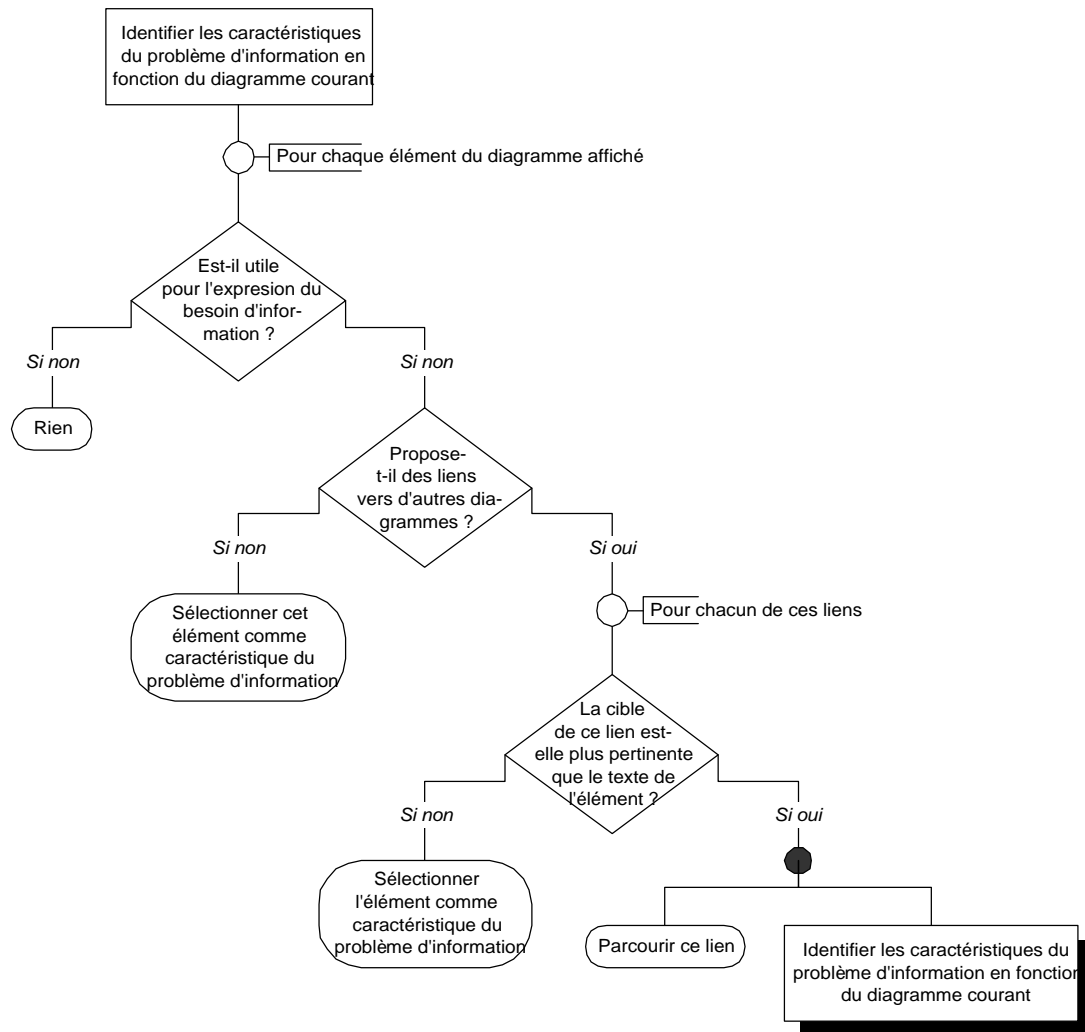


Figure 30. Détails de la tâche récursive de navigation hypermédia pour la sélection d'un élément significatif pour le problème d'information.

Ces diagrammes font apparaître d'une part le caractère itératif de la recherche d'information (déjà souligné au chapitre 2) et d'autre part le caractère « récursif »¹¹³ de la stratégie d'accès à l'information cumulative. Cette récursivité est due à la navigation hypermédia dans la collection de diagrammes de connaissances. Elle peut poser pour l'utilisateur un problème de surcharge cognitive, de désorientation ou d'oubli du but de la recherche¹¹⁴.

¹¹³ En informatique, la récursivité est traitée comme tout appel de fonctions en utilisant une pile. L'« environnement » (i.e. les variables locales et la position dans la fonction où a lieu l'appel) est empilé, la fonction appelée est traitée, et le retour à la fonction appelante s'effectue en dépilant l'environnement. La particularité de la récursivité est de permettre à la même fonction de s'appeler indéfiniment. Un programme récursif peut alors facilement générer des piles de tailles conséquentes, voire un dépassement de la taille maximale de la pile (« stack overflow »).

¹¹⁴ En effet, un utilisateur qui effectue une recherche d'information par navigation « empile » dans sa « mémoire de travail » (cf. [Rouet & Tricot 98]), les différentes informations qu'il a à stocker. La première d'entre elles est le problème d'information. C'est donc la plus lointaine des informations dont il a à se souvenir lorsqu'il réalise plusieurs étapes de navigation hypermédia.

Pour éviter cela, il s'agit de limiter la profondeur de la récursivité, et ainsi d'éviter les « débordements de pile ». Cela implique de réduire le nombre d'étapes nécessaires pour atteindre chaque diagramme de la collection. Trois solutions sont proposées ici. La première concerne le contenu des diagrammes. De fait, les relations sémantiques définies dans la méthode MKSM permettent d'envisager une grande partie des inférences qui peuvent être faites entre la tâche courante de l'utilisateur (i.e. celle qui a provoqué sa démarche d'accès à l'information) et le reste de la modélisation. C'est alors au chargé de projet d'identifier avec l'expert un maximum de ces relations, de sorte qu'elles apparaissent dans le résultat de la modélisation. Les deux autres solutions concernent les éléments d'interface et sont développées au paragraphe suivant.

Une autre solution possible au problème de désorientation est de permettre à l'utilisateur de changer de stratégie en cours de recherche s'il juge avoir atteint les limites de ses capacités de mémorisation. De plus, pour [Marchionini 95], le fait de changer de stratégie au cours de la recherche est courant et n'est pas nécessairement induit par une désorientation de l'utilisateur. Celui-ci peut par exemple décider de sélectionner certains éléments des diagrammes puis préciser les caractéristiques particulières de son problème qui ne se trouvent pas dans la modélisation. Les diagrammes ci-dessus font d'ailleurs apparaître que les différentes stratégies envisageables sont proches et partagent certaines tâches. C'est pourquoi l'interface ne doit pas cloisonner ces stratégies, mais permettre le passage de l'une à l'autre sans perdre le bénéfice acquis lors des étapes précédentes.

Dans ce cas, l'utilisateur peut compléter la description de son besoin d'information en fournissant des descripteurs des documents recherchés sans recourir à la navigation dans les diagrammes de connaissances. Le LMCE lui permet d'articuler ces descripteurs de la même manière que les éléments des diagrammes et de conserver les éléments sélectionnés pendant les étapes de navigation antérieures. Le LMCE effectue ensuite de la même façon la formalisation de la requête pour lui. Pour cela, des éléments d'interface spécifiques sont mis à sa disposition.

4.2.4.2.2. Éléments d'interface

Plusieurs types d'éléments d'interface sont nécessaires à la réalisation des tâches d'accès à l'information. Il s'agit d'une part du bandeau de titre, des diagrammes de connaissances, des cartes conceptuelles et du multiplexeur de liens, décrits au paragraphe précédent, et d'autre

part d'un formulaire d'interrogation spécifique, baptisé *caddie à concepts*, et de raccourcis vers les diagrammes de connaissances.

4.2.4.2.2.1. Bandeau de titre, diagrammes et cartes conceptuelles

Hormis le fait que l'activité générale indiquée sur le bandeau de titre est ici la recherche d'information cumulative, cet élément ne change pas par rapport au cas précédent. Il en va de même pour les diagrammes de connaissances.

Le recours aux cartes conceptuelles permet de limiter les risques de surcharge cognitive en diminuant le nombre d'étapes nécessaires pour accéder à un diagramme donné. C'est en outre un moyen commode pour faciliter la navigation. Leur intérêt pour la recherche d'information est de permettre une classification par formalismes. Ainsi, un utilisateur qui cherche un type de document sait qu'il le trouvera sous forme de flux dans un diagramme d'activité, et celui qui cherche un concept le trouvera dans un diagramme du même nom. Les titres des diagrammes étant mentionnés dans chaque carte, l'utilisateur n'a alors aucun mal à trouver l'élément dans le diagramme recherché.

Par ailleurs, d'après [Roger & al. 98], la distance entre les tâches de navigation dans les hypermédias destinés à l'apprentissage et à l'accès à l'information est faible. Cela est dû notamment au fait que l'accès à l'information est une des composantes des processus d'apprentissage. Les distinctions que ces auteurs relèvent dans la littérature ne portent pas sur les points abordés au paragraphe précédent. En conséquence, nous considérons ces trois points comme valables pour l'accès à l'information utilisant la navigation dans les diagrammes de connaissances. La présentation des cartes conceptuelles, la particularisation des sources des liens et l'explicitation de la sémantique des liens sont réalisées dans ce cas comme dans le cas précédent.

4.2.4.2.2.2. Le multiplexeur de liens

L'utilisateur du LMCE se sert de la navigation hypermédia pour accéder aux diagrammes. Même si elle est une sous-tâche de la recherche d'information, nous considérons que la tâche de navigation est ici identique à celle décrite dans le paragraphe précédent. Cette tâche est donc réalisée par l'intermédiaire du multiplexeur de liens présenté plus haut.

Cependant, le but de la navigation est ici la sélection d'éléments des diagrammes pour la recherche d'information. Afin de permettre cette sélection, nous avons ajouté, comme cela est fait dans le prototype de Jean-Marc Charlot, un item dans ce multiplexeur de liens. Cet item

permet la sélection de l'élément pour la stratégie d'accès à l'information cumulative. En « cliquant » sur cet item, l'utilisateur provoque l'envoi des caractéristiques de l'élément du diagramme vers le caddie à concepts décrit ci-dessus.

4.2.4.2.2.3. Le caddie à concepts

Pour la stratégie d'accès à l'information cumulative, l'utilisateur du LMCE a à sa disposition un formulaire HTML qui lui permet de collecter des éléments des diagrammes pendant sa navigation. La sélection d'un élément d'un diagramme provoque l'envoi vers ce caddie à concepts du texte de cet élément ainsi que de sa valeur de reformulation pour la recherche d'information¹¹⁵. Le texte de cet élément s'affiche dans le premier champ libre s'il en existe un. Sinon, il s'affiche dans le dernier champ et décale tous les autres champs, de sorte que l'élément le plus ancien est perdu.

Dans le cas où l'utilisateur renseigne ces champs par un autre moyen que la sélection d'éléments des diagrammes, il peut les taper au clavier, mais également utiliser les mécanismes de copier-coller ou de glisser-déplacer à partir du texte d'un autre document, qui sont gérés par le navigateur HTML. Ceci à condition bien entendu que le format du document d'origine permette ces mécanismes (i.e. que le logiciel de visualisation gère les liens OLE).

Le besoin d'information est exprimé par l'utilisateur en articulant les textes sélectionnés dans ce formulaire. Au moyen d'opérateurs logiques (ET, OU, opérateur de proximité¹¹⁶...), l'utilisateur peut construire une articulation sémantique de tous ces concepts qui constitue l'expression de son besoin d'information. Par défaut, le système relie tous ces éléments par des ET. Pour cela, l'utilisateur a à sa disposition des « radio buttons » qui lui permettent de sélectionner un opérateur de liaison entre le contenu d'un champ et celui du précédent.

En parallèle, la requête est construite en s'appuyant sur les reformulations des éléments des diagrammes ou sur le texte entré par l'utilisateur et en respectant le langage d'interrogation. L'utilisateur a également accès à la formulation de la requête correspondant à l'expression de son besoin d'information dans le langage d'interrogation du moteur de recherche. Cette requête est visualisable et modifiable dans un champ spécifique du caddie à concepts.

¹¹⁵ Le caddie à concepts présente à l'utilisateur les textes de ces éléments, mais construit la requête à partir de leur reformulation.

¹¹⁶ Opérateur permettant de spécifier une distance maximale entre deux termes recherchés (« NEAR » dans le langage d'interrogation VDK de Verity).

Lorsque l'utilisateur estime avoir correctement exprimé (ou formulé) son besoin d'information (ou sa requête, s'il a choisi de modifier celle-ci dans le champ correspondant), il la soumet au système d'interrogation qui fournit les réponses correspondantes. Pour cela, il clique sur un bouton du formulaire qui soumet la requête au système.

Nous avons mis au point un tel formulaire HTML, qui a été effacé à la fin de la thèse au CEA. La figure ci-dessous montre une version antérieure de ce formulaire, qui mêle interrogation par champs et stratégie d'accès à l'information cumulative. Elle permettait de sélectionner les index à partir des champs et de construire une recherche à partir du texte des champs. Il s'est d'ailleurs avéré que le premier champ ne peut être associé à un opérateur booléen binaire (ET ou OU) mais uniquement à l'opérateur unaire NON.

Le format du formulaire ci-dessous n'est pas compatible avec le modèle d'écran retenu pour la disposition des éléments d'interface du LMCE (voir paragraphe suivant). De plus, il peut créer une confusion entre les différentes stratégies d'interrogation proposées¹¹⁷. Nous le présentons cependant ici car il permet d'illustrer la différenciation faite par le prototype entre les critères d'indexation de fonds et de forme. En outre, le haut de ce formulaire permet également d'illustrer le formulaire spécifique au domaine pour l'interrogation par champs.

RECHERCHER	
un document du type	Note SEC/T
datant de	1995
dont l'auteur est	inconnu
dans le domaine	Expériences critiques
contenant les termes ou expressions :	
<input checked="" type="radio"/> et <input type="radio"/> ou <input type="radio"/> non	U235 enrichi à 4.75%
<input checked="" type="radio"/> et <input type="radio"/> ou <input type="radio"/> non	acide nitrique
<input type="radio"/> et <input type="radio"/> ou <input checked="" type="radio"/> non	approche par montée d'eau
<input checked="" type="radio"/> et <input type="radio"/> ou <input type="radio"/> non	

Figure 31. La première version du caddie à concepts. Les textes figurant dans les champs sont issus d'éléments des diagrammes du projet-pilote.

Sur ce formulaire comme sur les suivants, nous avons fait le choix de permettre l'articulation des éléments de la requête en les insérant dans une phrase en langage naturel. Nous escomptons qu'ainsi, l'utilisateur non expérimenté en matière de recherche documentaire est

¹¹⁷ Une telle confusion peut par exemple se traduire par une redondance du choix des critères d'indexation entre les deux stratégies et engendrer le sentiment de « perte de contrôle de la situation » (cf. [Bastien & al. 98] p.153).

capable d'articuler ces éléments lors de la phase de construction de la requête¹¹⁸. Nous avons réalisé un autre formulaire, que nous considérons comme valide pour le LMCE. Ce formulaire est présenté ci-dessous et est celui décrit dans le reste de ce paragraphe. Les caractéristiques graphiques (polices, couleurs) n'ont pas encore été déterminées pour ce formulaire.

Figure 32. Le caddie à concepts tel qu'il est proposé dans le LMCE.

Sur cette figure, nous constatons une redondance entre le texte d'introduction du formulaire (« Chercher les documents ») et celui du bouton d'appel de la procédure de recherche (« Rechercher »). Nous avons choisi cette redondance afin de ne pas imposer à l'utilisateur de « revenir au début du formulaire pour aller à la fin de sa requête ». Ce choix a pour effet de charger la barre de boutons. C'est pour cela que le texte du bouton, supposé plus important, est présenté en gras. Nous reconnaissons donc que ce choix est discutable. Toutefois, il n'est pas discuté ici et nous comptons sur l'évaluation du prototype pour le valider ou l'infirmier. Dans ce dernier cas, il est techniquement simple d'insérer le bouton d'appel de la recherche à la place et avec le texte de l'introduction du formulaire.

Techniquement, le caddie à concepts est réalisé sous forme d'un cadre HTML dans lequel les autres cadres de la page peuvent envoyer des données via le DOM Microsoft®. Il est lui-même composé de plusieurs cadres, ce qui permet de redimensionner la fenêtre en conservant l'accès aux boutons. Une fenêtre redimensionnée est montrée dans la figure suivante. De la même façon, les boutons interagissent avec les champs par l'intermédiaire du DOM Microsoft®.

¹¹⁸ C'est par exemple également le parti pris par les concepteurs du site « Hcibib.org » de consultation d'une base de données en ligne d'ergonomie des logiciels. Ce parti pris est défendu dans la FAQ de ce site (<http://www.hcibib.org>).

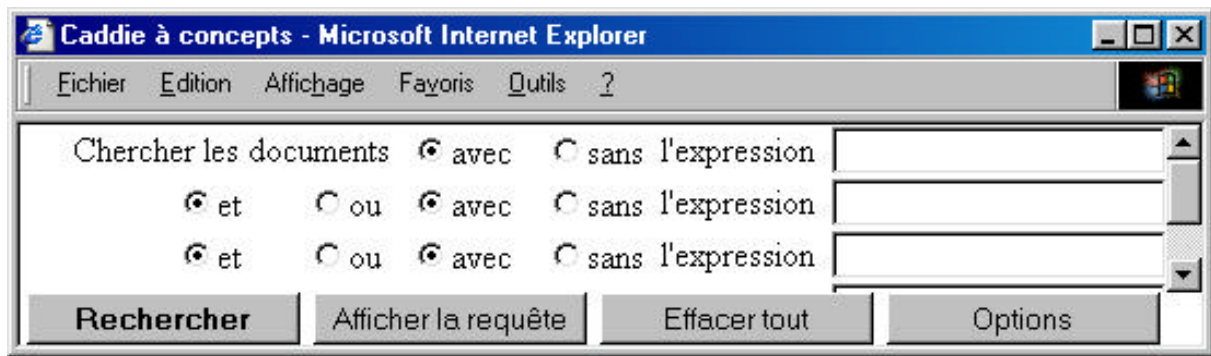


Figure 33. le caddie à concepts redimensionné.

Ces fonctions dynamiques sont réalisées par l'intermédiaire de scripts en langage Jscript (i.e. la version de JavaScript de Microsoft Internet Explorer®). L'exemple de fonction Jscript donné ci-dessous permet d'afficher ou de faire disparaître le champ texte présentant la formulation de la requête construite à partir des éléments des autres champs. L'action de cette fonction est de faire apparaître ou disparaître une table contenant ce champ texte et le bouton d'actualisation situé à droite de ce champ, comme cela est illustré dans la figure suivante.

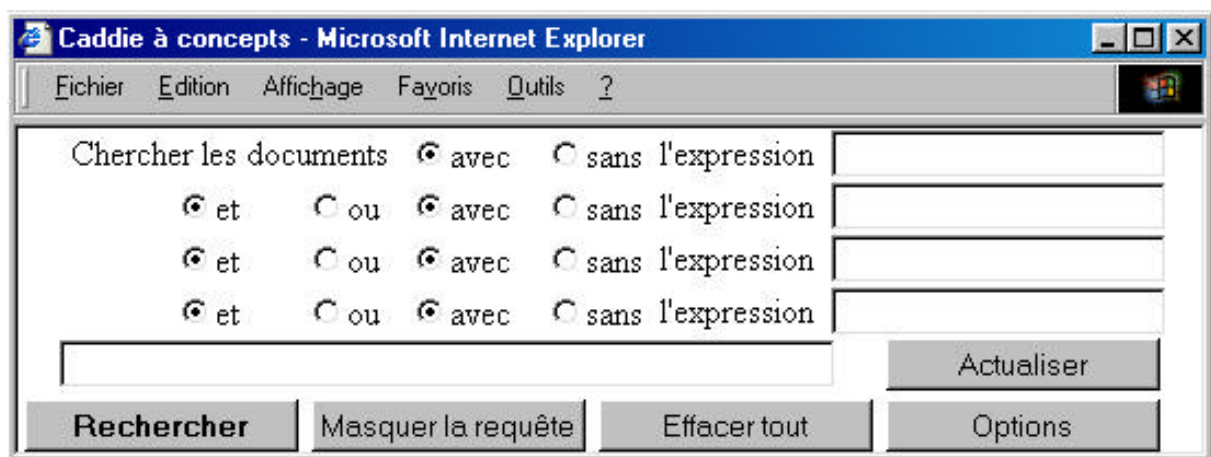


Figure 34. Le caddie à concepts avec l'affichage de la requête.

```
function button2_onclick()
{
    if (document.parentWindow.parent.window.document.frames("champs").window.
        document.all.requete.style.visibility=="hidden")
    {
        document.parentWindow.parent.window.document.frames("champs").window.
            document.all.requete.style.visibility="visible";
        document.parentWindow.parent.window.document.frames("champs").window.
            document.all.requete.style.height="24px";
        document.parentWindow.parent.window.document.frames("champs").window.
            document.all.requete.all.text5.style.height="22px";
        document.parentWindow.parent.window.document.frames("champs").window.
            document.all.requete.all.button1.style.height="24px";
        document.parentWindow.parent.window.document.frames("champs").window.
            document.all.hr1.style.visibility="hidden";
        button2.value="Masquer la requête";
    }
}
```

```

else
{
    document.parentWindow.parent.window.document.frames("champs").window.
        document.all.requete.style.visibility="hidden";
    document.parentWindow.parent.window.document.frames("champs").window.
        document.all.requete.style.height="0px";
    document.parentWindow.parent.window.document.frames("champs").window.
        document.all.requete.all.text5.style.height="0px";
    document.parentWindow.parent.window.document.frames("champs").window.
        document.all.requete.all.button1.style.height="0px";
    document.parentWindow.parent.window.document.frames("champs").window.
        document.all.hr1.style.visibility="visible";
    button2.value="Afficher la requête";
}
}

```

Le bouton correspondant est défini dans une table du corps de la page HTML par l'élément :

```

<INPUT id="button2" name="button2" type=button value="Afficher la requête"
style="HEIGHT: 24px; WIDTH: 125px" LANGUAGE="javascript" onclick="return
button2_onclick()">

```

Le bouton « Rechercher » est le déclencheur du contrôle ActiveX d'interrogation du moteur de recherche. Ce contrôle identifie les éventuels index liés aux critères de fond en fonction des termes contenus dans la requête. Il sélectionne les sous-index du profil de l'utilisateur correspondants et élimine les termes correspondants de la requête. Pour cela, il tient compte des opérateurs logiques entre les concepts de la requête.

Pour illustrer cela dans le contexte de notre projet-pilote, nous prenons l'exemple d'un utilisateur qui s'intéresse au sous-domaine des transports de matières fissiles et désire un filtrage restreint. Si sa requête est « PU AND "arrangement spécial" »¹¹⁹, les index sélectionnés seront des sous-index de toutes les catégories de documents indexées, parmi ces sous-index, le mode de filtrage ne conserve que les sous-index associés au sous-domaine des transports de matières fissiles. Enfin, le contrôle ActiveX reconnaît une catégorie de matière fissile (critère de fond) dans la requête formulée et sélectionne donc, parmi les catégories restantes, les index portant sur le type de matière fissile « plutonium ». La requête envoyée au moteur de recherche est alors uniquement « "arrangement spécial" ».

4.2.4.2.2.4. Les raccourcis

Toujours dans le but de diminuer le nombre de « clicks » nécessaires pour accéder à un diagramme, nous proposons l'emploi de raccourcis vers les types de diagrammes les plus souvent accédés ou ceux reconnus par les experts comme les plus importants pour l'accès à l'information dans leurs sous-domaines respectifs. Ces raccourcis sont fournis soit sous la

¹¹⁹ Note : en majuscules, PU désigne un type de matières fissiles contenant essentiellement du plutonium. Un arrangement spécial est une autorisation de transport de matières fissiles ponctuelle.

forme de « favoris », soit sous celle de « liens » du navigateur. L'avantage de cette dernière solution est qu'ils restent alors toujours visibles, et donc accessibles d'un seul « click ». Cette solution est de plus facile à mettre en place et aisément personnalisable par les utilisateurs.

4.2.4.2.3. *L'interrogation directe*

La stratégie d'interrogation directe du fonds documentaire s'effectue avec le LMCE comme avec n'importe quel autre outil. L'analyse des tâches correspondant à cette stratégie n'est pas menée ici. L'utilisateur qui ne souhaite pas utiliser la navigation dans les diagrammes de connaissances a à sa disposition un formulaire d'interrogation simple pour les domaines de recherche et comportant plusieurs champs pour les domaines d'ingénierie. Dans ce dernier cas, les champs du formulaire sont du même type que ceux de la figure présentée plus haut.

Ce formulaire est relié au même contrôle ActiveX d'interrogation que pour la stratégie cumulative. Ainsi, l'utilisateur bénéficie du filtrage de l'information selon son profil quelle que soit la stratégie d'accès à l'information adoptée.

Le seul autre élément d'interface affiché ici est le bandeau de titre. Celui-ci indique pour cette activité la mention « Recherche d'information directe ».

4.2.4.3. Les autres fonctions du LMCE

Pour pouvoir fonctionner correctement, le LMCE doit présenter d'autres fonctions « accessoires », qui sont de permettre à un utilisateur de se créer un profil et de s'identifier. Faute de temps, nous n'avons pas accordé d'attention particulière à la réalisation de ces tâches. À terme il est donc nécessaire de revenir sur ces points. Nous avons donc choisi des solutions « classiques ».

La définition du profil se fait dans un formulaire accessible dans une page HTML séparée du reste de l'interface. Les champs sont renseignés à l'aide de « radio buttons », de « combo boxes » ou de « text fields » selon les types de données à indiquer. Le fichier XML correspondant au profil est créé ou mis à jour par un script de la page.

La connexion se fait dans la « page de garde » du LMCE, en demandant à l'utilisateur son nom et mot de passe dans des champs textes spécifiques (le mot de passe ne s'affiche pas). L'interaction dans une page HTML plutôt que dans une boîte de dialogue est justifiée par le fait que pour les pages web, le navigateur Microsoft Internet Explorer® permet à l'utilisateur de mémoriser le contenu de ces champs et de lui éviter de les retaper à chaque connexion. Ce formulaire comporte en plus la possibilité de définir un nouveau profil, sous la forme d'un

lien vers le formulaire précédent. Le refus de connexion (nom de l'utilisateur et mot de passe ne correspondant pas à un document XML de la base des profils) est signifié par l'affichage d'un message d'erreur au-dessus du formulaire de connexion. L'acceptation de la connexion provoque l'affichage de la première page à visualiser pour l'activité générique d'utilisation du LMCE pour le profil considéré (i.e. navigation, interrogation cumulative, interrogation directe).

Les fonctionnalités de présentation des résultats d'une requête et de visualisation des documents ne sont pas abordées ici. En effet, jusqu'ici, nous n'avons aucun contrôle sur le fonctionnement du moteur de recherche après l'envoi d'une requête. Celui-ci (Acrobat Search®) présente les résultats de la recherche dans une boîte de dialogue si plusieurs documents sont trouvés ou en ouvre directement un s'il est la seule réponse trouvée à la requête. L'ergonomie de cet outil est discutable, mais le LMCE en est pour l'instant tributaire. Dans l'optique d'une indexation et d'une recherche par le moteur de recherche SPIRIT®, comme cela est envisagé¹²⁰, la présentation des résultats de la requête aurait lieu dans un cadre de l'interface du LMCE via l'interface W3 de ce moteur de recherche ([Fluhr & al. 97]). Il s'agirait alors d'homogénéiser cette interface avec celle du LMCE. Du fait de son caractère encore hypothétique, une telle homogénéisation n'est pas discutée dans ce travail, même si elle est importante pour l'ergonomie du LMCE.

De la même façon, la visualisation des documents ne peut se faire que dans une fenêtre séparée. Il est en effet impossible d'exécuter un plug-in Acrobat (ici, le plug-in de mise en surbrillance des mots retrouvés dans le document) dans une instance de ce logiciel incorporée à un navigateur HTML (cf. [Adobe 99]).

4.2.4.4. Les aspects globaux de l'interface

Les paragraphes ci-dessus proposent une démarche de conception des éléments de l'interface du LMCE fondée sur la description des tâches de l'utilisateur. Il s'agit ici de définir la mise en place de ces éléments de l'interface en fonction des tâches à réaliser.

Trois grands types d'activités sont identifiés : la navigation simple dans les diagrammes de connaissances, la recherche d'information grâce à la stratégie cumulative et les recherches d'informations utilisant des stratégies analytiques (interrogation directe ou par champs). Dans

¹²⁰ Nous avons d'ailleurs encadré un projet d'un groupe d'étudiants du DEA IST de l'INSTN (CEA de Saclay) sur ce thème. Ce projet a permis l'indexation de fonds pdf mais laissé en suspens la présentation des résultats

ce paragraphe, nous rappelons pour chacune de ces activités les éléments d'interface nécessaires et en proposons une disposition à l'écran.

Pour cela, nous nous appuyons sur le modèle d'écran proposé par [Asse 85] (via [Kolski 97] p. 221) pour définir l'aspect général de l'interface. Ce modèle divise l'écran en trois zones horizontales comme illustré dans la figure ci-dessous. La zone supérieure contient les informations générales telles que le titre de la page. La zone centrale contient le contenu signifiant de la page. Cette zone permet la disposition des informations qui conduisent l'utilisateur à interagir avec le système. Enfin, la zone inférieure est utilisée pour permettre à l'utilisateur de contrôler le fonctionnement du système. Dans la suite, ces trois zones sont appelées espace titre, espace d'échange d'informations et espace de commandes utilisateur, selon une terminologie proche de celle de l'auteur de ce modèle.

Le choix de ce modèle est justifié par le fait qu'il est adopté – plus ou moins consciemment – par les concepteurs de nombreux sites web et applications. L'intérêt que présente ce modèle pour la conception du LMCE est d'homogénéiser les différents aspects de l'interface. Quelle que soit la tâche qu'il est en train d'accomplir, l'utilisateur est à tout moment capable d'identifier la portée des informations qui sont présentes à l'écran.

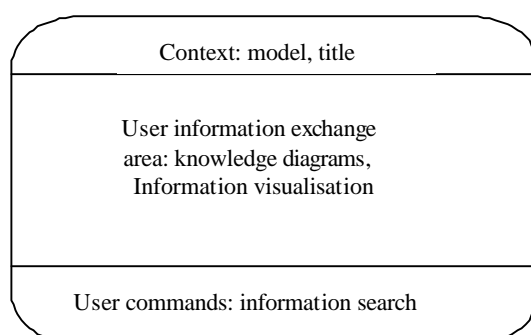


Figure 35. Le modèle d'écran d'A. Asse.

4.2.4.4.1. Navigation dans les diagrammes de connaissances

Comme nous l'avons vu, la navigation simple s'adresse aux utilisateurs novices. Quatre éléments d'interface sont nécessaires pour cette activité : le bannière de titre, les diagrammes de connaissances, les cartes conceptuelles et les multiplexeurs de liens.

4.2.4.4.2. Espace titre

La première information à fournir à l'utilisateur pour l'aider à comprendre le fonctionnement du LCH est le type d'information qu'il est en train de visualiser. C'est pourquoi il nous paraît utile de faire figurer dans cet espace le bandeau de titre.

4.2.4.4.3. Espace d'échange d'informations

La navigation est l'activité principale de l'utilisateur dans le cas qui nous occupe. C'est donc par cette activité qu'est réalisé l'échange d'informations. Nous plaçons donc dans cette zone les éléments d'interface destinés à cette activité, c'est-à-dire les diagrammes, les multiplexeurs de liens et les cartes conceptuelles. Les multiplexeurs de liens étant relatifs aux éléments des diagrammes, nous avons vu qu'ils apparaissent par-dessus les diagrammes, à proximité des éléments auxquels ils se rapportent.

Les diagrammes et les cartes conceptuelles étant liés, il semble logique de les représenter à l'écran dans un espace commun. Il paraît en outre logique que ces deux éléments soient dans une position centrale de l'écran, car ils représentent l'information la plus importante pour l'utilisateur. Nous avons choisi d'aligner ces deux éléments horizontalement, avec la carte à gauche des diagrammes, comme cela est présenté sur la plupart des sites proposant une carte de navigation. C'est en outre la disposition proposée dans le prototype de Jean-Marc Charlot.

4.2.4.4.4. Espace de commandes utilisateurs

Dans le cas de l'utilisation du LMCE pour la formation, nous n'avons pas d'autre élément à faire figurer dans cette zone. Plutôt que de laisser un « vide à l'écran, nous préférons supprimer cette zone, et ainsi permettre l'utilisation de la place vacante pour l'affichage des éléments des autres zones. Cela est d'ailleurs fait automatiquement par le navigateur, lorsqu'un élément « FRAME » est vide et que les autres ne peuvent pas être entièrement affichés (i.e. il existe une barre de défilement dans l'un d'eux)¹²¹.

À terme, il peut être intéressant d'envisager la présentation dans cet espace d'aides supplémentaires à la navigation. Cela a été fait au début de la thèse, mais nous n'avons pas persisté dans cette voie et avons préféré nous concentrer sur les tâches d'accès à l'information. Il a notamment été envisagé d'explicitier les liens dont le diagramme courant est à la source. Cette information supplémentaire se présente sous forme de liste et est supposée

¹²¹ À condition que l'espace alloué à chaque cadre soit défini en relatif (i.e. en pourcentage de l'espace total) et non en absolu (i.e. en pixels).

être lue après le diagramme. C'est pourquoi il nous semble logique de la placer en dessous de l'espace réservé à la carte et aux diagrammes. La figure suivante montre une ancienne version (statique) du prototype proposant cette fonctionnalité.

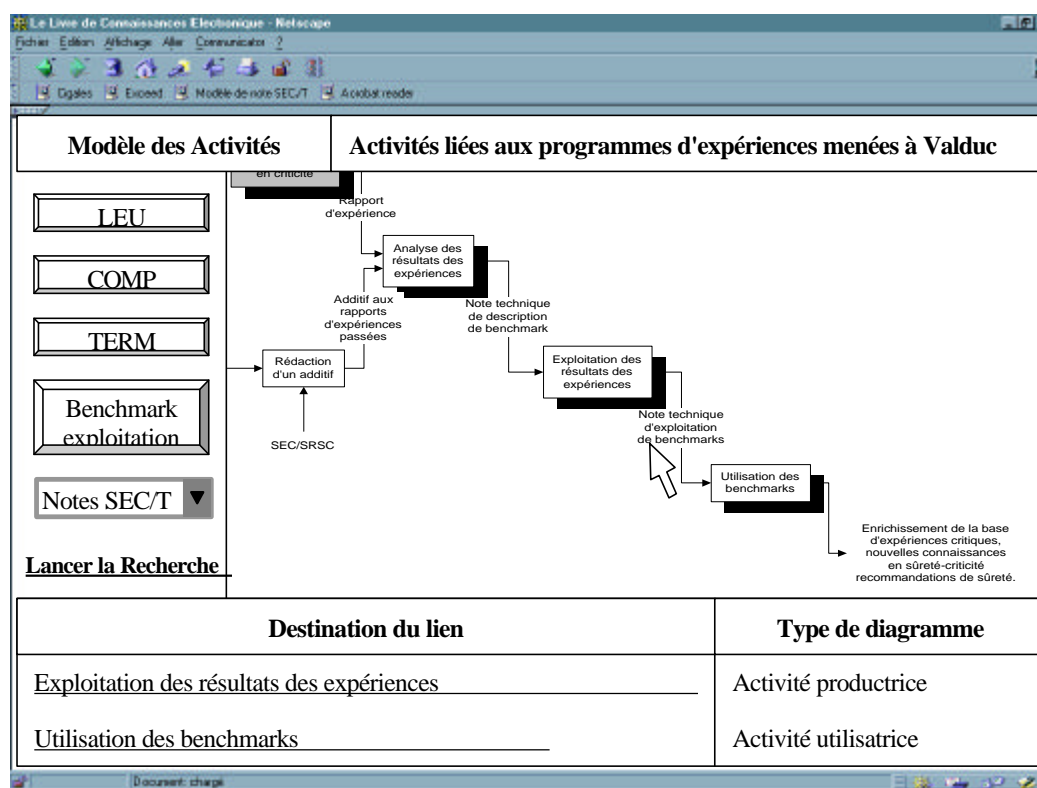


Figure 36. Un prototype d'interface du LMCE pour les utilisateurs novices.

4.2.4.4.5. Accès à l'information utilisant la stratégie cumulative

Pour cette activité, nous avons vu que les quatre éléments d'interface mentionnés au paragraphe précédent sont conservés. La navigation étant toujours une tâche importante de la stratégie d'accès à l'information cumulative, ces éléments conservent les places indiquées ci-dessus. En plus de ces quatre éléments, nous avons ici besoin de positionner le caddie à concepts et les raccourcis.

Le caddie à concepts s'insère dans l'espace de commandes de l'utilisateur. En effet, la stratégie cumulative se termine par l'articulation des éléments sélectionnés lors de la navigation et par l'envoi de la requête. Ces deux dernières tâches ont lieu dans le caddie à concepts. C'est pourquoi il nous paraît logique d'afficher ce dernier élément en dessous de ceux relatifs à la navigation.

Les raccourcis sont gérés par le navigateur et ne prennent pas place dans la page HTML visualisée.

4.2.4.4.6. Accès à l'information par interrogation directe

Cette stratégie d'accès à l'information ne nécessite que le bandeau de titre et un formulaire spécifique d'interrogation. Le bandeau est affiché dans l'espace titre et le formulaire d'interrogation occupe le reste de la page.

4.2.4.4.7. Autres fonctions

Les autres fonctionnalités sont réalisées dans des pages HTML homogènes avec celles présentées ci-dessus. Elles comportent notamment un bandeau de titre spécifiant l'activité concernée (définition/modification d'un profil ou connexion d'un utilisateur au LMCE). Dans l'espace d'échange d'informations se trouve le formulaire concerné par l'activité en cours. L'espace de commandes utilisateur permet d'afficher au-dessous les boutons de validation, d'effacement et de sortie sans validation. Cette sortie est dirigée :

vers la dernière page visualisée (permettant la navigation ou l'accès à l'information) pour la modification du profil d'un utilisateur déjà connecté,

vers le formulaire de connexion pour la création ou la modification d'un profil lorsque l'utilisateur ne s'est pas déjà connecté et

vers le formulaire de création d'un profil à partir du formulaire de connexion.

L'idée ici est d'assurer l'homogénéité des différents écrans en présentant ces fonctionnalités avec les mêmes feuilles de style (CSS et XSL).

Dans l'optique où la fonctionnalité d'indexation par SPIRIT® du fonds documentaire serait possible, la visualisation des résultats des requêtes se présenterait alors de la même manière. Il est même possible d'envisager le même type de présentation pour la visualisation des documents, si Adobe arrive à résoudre les problèmes techniques qui empêchent actuellement les plug-ins Acrobat® de fonctionner lorsque l'application est incluse dans une autre fenêtre. Dans ce dernier cas, l'interface du LMCE comporterait un bandeau de titre placé en haut de l'écran comportant par exemple la mention « Consultation des documents » et éventuellement le titre du document, si celui-ci est renseigné. Le cadre central contiendrait la fenêtre Acrobat permettant la visualisation des documents. L'espace de contrôle situé en dessous proposerait le retour aux diagrammes de connaissances, voire un formulaire d'interrogation permettant la recherche à partir d'éléments des documents¹²².

¹²² Nous avons testé cette possibilité sur la visualisation simple de documents et il apparaît que le glisser-déplacer et le copier-coller à partir de sélections de textes dans un document pdf visualisé dans un cadre du navigateur sont pris en charge par Acrobat® et Internet Explorer®.

4.2.4.5. Critères ergonomiques

Les paragraphes ci-dessus présentent notre démarche de conception de l'interface du LMCE. Notre préoccupation est de favoriser l'utilisabilité de cette interface. Dans ce paragraphe, nous passons en revue les critères de conception ergonomiques indiqués par [Bastien & al. 98]. L'objectif est ici d'indiquer les critères qui ont été pris en compte dans cette conception et d'identifier ceux pour lesquels l'interface peut être améliorée. La discussion de ces critères, même si elle figure a posteriori dans ce manuscrit, fait apparaître la prise en compte de ces critères pendant le processus de conception. Pour beaucoup d'entre eux, la discussion fait d'ailleurs référence aux chapitres précédents.

Nous sommes par ailleurs conscient que pour être avérée, l'utilisabilité du prototype doit être validée. Cela n'a pas été fait avant la rédaction de ce manuscrit car le « prototype final » du LMCE tel qu'il est présenté ici est en cours de développement. Nous escomptons avoir des résultats complémentaires sur ce point à présenter lors de la soutenance.

Incitation : pour Christian Bastien & al., ce critère est l'un des plus importants à prendre en compte pour la conception d'un hypermédia. C'est d'ailleurs celui auquel nous nous sommes le plus intéressé pour la conception du LMCE. La plupart des points figurant dans ce critère (comme le typage des liens, la redondance avec le diagramme cible ou les cartes conceptuelles) ont été traités plus haut. Seule la présentation des aides à la navigation dans chaque page n'est pas réalisée. Il nous semble en effet qu'une telle insertion de pointeurs de navigation va à l'encontre du critère d'homogénéité pour les diagrammes de connaissances. Cependant, la présence des cartes conceptuelles à côté des diagrammes compense à notre sens cette lacune.

Concernant l'incitation proprement dite (c'est-à-dire les mécanismes qui permettent de faire comprendre à l'utilisateur les actions que le système attend de lui), nous distinguons les tâches usuelles (remplissage du formulaire de saisie du profil, connexion, envoi de la requête et sélection d'un document à visualiser dans la liste d'items retrouvés) et les tâches spécifiques au LMCE. Pour ces dernières, les éléments d'interface mentionnés plus haut ne sont pas familiers à l'utilisateur. En particulier, la sélection des liens de navigation ou le typage des éléments des diagrammes de connaissances sont spécifiques au LMCE. De la même manière, la stratégie cumulative d'accès à l'information peut paraître déroutante à un utilisateur non averti. Cet outil nécessite donc une prise en main de l'utilisateur. Cela peut être fait par des séances de présentation/formation sur cet outil, organisées dans l'organisation modélisée ou

par un mode d'emploi approprié. Toutefois, la mise en place d'un système d'aide doit à terme être envisagée.

Groupement/distinction entre items : la distinction entre les items se fait principalement par leur format. Les éléments d'interface présentés dans cette partie sont tous facilement identifiables. Les éléments de type cadres HTML (diagrammes, cartes conceptuelles et caddie à concepts) ont en plus une localisation prédéfinie. Ces deux critères sont donc valables pour ces éléments.

L'emploi de cadres pour les pages web est critiqué dans ce critère. Ils posent des problèmes de définition des cibles des liens, d'impression, d'indexation et de temps de chargement. L'emploi de l'attribut « TARGET » pour le choix d'un cadre d'affichage de la cible d'un lien est recommandé. Cet attribut, déjà utilisé dans le prototype de Jean-Marc Charlot, permet à la fois de conserver l'homogénéité des pages et d'éviter les « rechargements » inutiles. L'action des boutons « Back » et « Forward » est discutée plus loin. Par ailleurs, l'interface du LMCE (hormis le fonds documentaire, qui est géré par une autre application) n'est a priori pas destinée à être indexée par un moteur de recherche sur le web.

Concernant l'impression des pages du LMCE, l'impression d'un diagramme isolé peut avoir du sens, il convient de permettre cette impression et de gérer les paramètres satisfaisants pour ce type de contenu. Cette fonctionnalité n'est pas traitée dans ce travail et nous nous limitons à proposer des pistes à explorer pour permettre cette impression. Cela est par exemple réalisable grâce à l'interface « IoleCommandTarget » et à la méthode « QueryStatusWB » du navigateur qui permettent de capturer et de spécifier un comportement pour l'impression. Il est alors possible de modifier les paramètres de l'impression (passage en mode « paysage » pour l'impression d'un diagramme, voire d'imprimer un autre document que celui affiché à l'écran. L'approche DVP permet en effet de composer dynamiquement un document HTML constitué de l'image exportée par Visio du diagramme, associée à une feuille de style appropriée (i.e. spécifiant par exemple uniquement le titre et le formalisme du diagramme). La question se pose alors d'évaluer la pertinence d'un item « Imprimer ce diagramme » parmi les options du multiplexeur de liens.

Feedback immédiat : chaque action de l'utilisateur provoque une réaction du système. Nous avons essayé de dresser une liste exhaustive de ces interactions dans les paragraphes précédents. De plus, les choix techniques (HTML dynamique et contrôles ActiveX) privilégient l'exécution du LMCE sur la machine cliente plutôt que sur un serveur. Du fait de

la relative simplicité des opérations à réaliser, le temps de réponse du système en est ainsi diminué.

Lisibilité : nous avons essayé de réaliser des éléments d'interface clairs et lisibles. Cependant, les différents points abordés dans ce critère ne sont pas passés en revue ici. Des progrès dans le sens de la lisibilité des éléments d'interface sont sûrement encore possibles. Par ailleurs, nous reconnaissons que de gros efforts doivent être faits au niveau de la mise sur écran des diagrammes de connaissances.

Concision : les éléments d'interface présentés ci-dessus satisfont aux deux critères de concision, à l'exception des diagrammes et des cartes conceptuelles. Pour ces éléments, nous sommes tributaires de la modélisation des connaissances réalisée. Par définition, le domaine modélisé est complexe et sa description nécessite un effort de compréhension – et donc une charge cognitive non négligeable – pour l'utilisateur. Le fait de disposer de plusieurs cartes conceptuelles spécifiques aux formalismes MKSM permet de réduire la taille de celles-ci et va donc dans le sens de la concision de l'interface.

Actions minimales : ce critère se justifie surtout pour les utilisateurs experts qui utilisent le LMCE pour l'accès à l'information. Il est pris en compte par exemple par l'utilisation des raccourcis ou des cartes conceptuelles. En revanche, dans le cas d'utilisateurs novices ou apprenants, nous avons privilégié la compréhension du domaine à la rapidité de la navigation. En particulier, le multiplexeur de liens oblige l'utilisateur à faire deux « clicks » pour atteindre la page suivante. Nous pensons en effet que dans un but d'apprentissage ou d'accès à l'information pour l'apprentissage, il est préférable de privilégier les critères de guidage et d'actions explicites à celui d'actions minimales.

Les problèmes de dimensionnement et de défilement des cadres ne sont pas traités ici. Les diagrammes Visio exportés en HTML sont par exemple souvent illisibles s'ils sont réduits à la taille d'un écran 1024x768. L'utilisation de différents cadres diminue encore l'espace de visualisation octroyé à ces diagrammes. Des solutions comme les loupes ou les « fisheye views » peuvent être étudiées.

Densité informationnelle : de la même façon, la majorité des éléments d'interface sont assez simples, à l'exception des diagrammes de connaissances. Pour certains diagrammes, l'information présentée est nécessairement dense. Cependant, cette information est jugée unitaire par les experts et à ce titre, le chargé de projet ne peut la morceler. De plus, le morcellement de l'information conduit à un ajout de liens qui n'est pas toujours bénéfique à la

concision de l'ensemble. Le critère concernant le nombre de liens par page est respecté, si l'on fait omission des cartes conceptuelles, dont le but est de proposer l'accès aux diagrammes. De même, les boutons radio n'ont pas plus de trois options.

Action explicite : ce critère est respecté pour toutes les actions de l'utilisateur, à l'exception d'une action. Nous avons vu que lors du parcours d'un lien inter-modèles, les cartes conceptuelles sont modifiées pour faire référence au formalisme du nouveau diagramme affiché. Ce choix est justifié plus haut.

Contrôle utilisateur : la plupart des tâches sont gérées par le navigateur. Dans ce cas, l'utilisateur a « la main » et peut interrompre un chargement en cours à tout moment. Toutefois, les contenus à charger sont peu volumineux et sont disponibles en Intranet. C'est pourquoi la plupart de ces chargements se font de manière quasi-instantanée. En revanche, la recherche d'information par le plug-in Acrobat est une tâche non instantanée sur laquelle l'utilisateur n'a aucun contrôle. Cependant, elle s'effectue à l'extérieur du navigateur, ce qui permet à l'utilisateur de continuer à interagir avec le LMCE pendant cette tâche. Par ailleurs, un problème qui a déjà été soulevé dans le prototype de Jean-Marc Charlot est la gestion des boutons « Back » et « Forward ». En effet, la navigation entre les diagrammes se fait en deux étapes : l'affichage du diagramme cible du lien sélectionné et la synchronisation des cartes conceptuelles avec ce nouveau diagramme. Chaque étape de navigation pour l'utilisateur est donc séparée de deux pages écran pour le navigateur. Cela doit donc être pris en compte au niveau de la gestion des boutons de la barre de navigation, comme cela est fait sur d'autres sites. Nous n'avons pas traité ce point dans ce travail.

Adaptabilité : les deux sous-critères de l'adaptabilité (flexibilité et prise en compte de l'expérience de l'utilisateur) sont pris en compte et ont été précédemment discutés.

Gestion des erreurs : trois types d'erreurs sont traités ici. L'ensemble des liens proposés étant interne au LMCE, l'erreur « Page not found » n'est pas censée intervenir dans l'utilisation de cet outil. L'erreur de syntaxe dans la formulation de la requête n'est pas prise en compte. Elle émane directement du moteur de recherche Acrobat Search, sur lequel nous n'avons aucun contrôle. Toutefois, elle ne concerne que les utilisateurs ayant choisi de formuler eux-même leurs requêtes. Ceux-ci sont capables de manier le langage d'interrogation du moteur de recherche et sont donc supposés comprendre ce message.

Le troisième « message d'erreur » envisagé est de ne pas obtenir de réponse à une requête informationnelle. La question du statut de ce message se pose : l'absence de réponse est-elle

une erreur du système, de l'utilisateur ou simplement une information significative sur le contenu du fonds documentaire ? Il nous semble qu'un élément de réponse à cette question réside dans la qualité de l'indexation et de la recherche documentaire. Idéalement, l'absence de réponse du moteur de recherche fournit l'information qu'il n'existe pas de document pertinent pour la requête. En pratique, les techniques d'indexation et de recherche étant rarement parfaites, il peut tout de même exister des réponses non trouvées par le système. Il s'agit alors d'une erreur du système. L'emploi d'une indexation conceptuelle et d'un moteur de recherche efficace (à terme, SPIRIT) est destiné à minimiser cette erreur. Enfin, il s'agit d'une erreur de l'utilisateur lorsque la requête n'est pas pertinente par rapport au problème d'information qu'elle est censée formaliser. Dans ce travail, nous nous efforçons de minimiser les erreurs du système et de l'utilisateur grâce à la formalisation de requêtes centrées sur l'activité de l'utilisateur. Le LMCE est conçu pour formuler des requêtes pertinentes à la fois par rapport au problème d'information de l'utilisateur et à l'indexation conceptuelle du système d'information.

Correction des erreurs : à part les moyens de correction des erreurs, proposés par le navigateur, ce critère n'a pas été pris en compte.

Homogénéité et cohérence : les informations disposées à l'écran sont homogènes et cohérentes entre elles à l'intérieur d'un même cadre. Nous avons montré précédemment que les différents cadres sont complémentaires. Une homogénéisation de ces cadres irait donc à l'encontre du critère de distinction entre les items. Plus généralement, l'approche DVP nous assure l'homogénéité de toutes les pages générées dynamiquement par le LMCE.

Signifiante des codes et dénominations : les sources des liens sont cohérentes avec leurs cibles, comme cela est précisé plus haut. Nous sommes cependant en contradiction avec la recommandation concernant l'emploi de termes techniques. Comme cela est précisé plus haut, le LMCE est dédié à un domaine complexe et fait référence à des notions complexes de ce domaine. Cependant, il nous semble que cette recommandation est spécifique à la conception de sites web supposés accessibles à tous. Elle ne s'applique pas à la conception de sites intranet, destinés à des utilisateurs identifiés.

Compatibilité : en matière de caractéristiques des utilisateurs, ce critère est traité par les profils utilisateurs. Concernant les matériels et logiciels de navigation, nous avons identifié plus haut l'environnement spécifique auquel est destiné le prototype discuté ici. La compatibilité de ce prototype avec d'autres types de matériels n'est pas traitée ici.

La discussion de ces critères fait apparaître leur prise en compte pendant le processus de conception. Toutefois, tous ces critères ne sont pas respectés. Le cas échéant, ce choix est justifié. Il peut être dû à des contraintes de temps ou aux caractéristiques particulières de l'outil en comparaison avec celles des sites web auxquels ces critères font référence. En particulier, la connaissance des utilisateurs et la finalité particulière du LMCE impliquent des choix de conception spécifiques. En tout état de cause, nous soulignons les points qui restent à améliorer.

Nous sommes donc conscient de l'imperfection de ce prototype. Nous estimons cependant avoir proposé une démarche de conception centrée sur l'utilisateur et orientée vers la maximisation de l'utilisabilité de l'interface. Même si cette démarche n'est pas conduite ici jusqu'à son terme, nous pensons que la structure du raisonnement permettant de le faire est présentée dans ce manuscrit et que la prise en compte des critères ergonomiques qui « manquent » est dès lors possible.

En d'autres termes, nous pensons que la construction du prototype présenté ici s'appuie sur des « bases ergonomiques saines » et espérons que les modifications qu'impliqueront nécessairement les résultats de la phase de tests ne nécessiteront pas de développements trop « lourds ».

4.2.5. *Spécifications du système*

Le paragraphe précédent clôt les spécifications de l'interface. À ce stade, nous avons une idée précise des différentes fonctionnalités du LMCE, des éléments d'interface nécessaires pour les réaliser, de l'organisation de ces éléments dans l'interface et du comportement de cette interface. Nous avons vu que certains points ne sont toutefois pas détaillés ici, comme les aspects graphiques du prototype. Nous connaissons également les détails des modèles sur lesquels s'appuie le LMCE. Il reste maintenant à aborder les spécifications du LMCE et la démarche opérationnelle de conception de cet outil. Pour cela, nous récapitulons l'ensemble des fonctionnalités énoncées dans les paragraphes précédents et proposons des spécifications techniques qui permettent d'aboutir à une architecture générique de cet outil. La démarche de conception est décrite dans un troisième paragraphe.

4.2.5.1. Fonctionnalités

Dans ce paragraphe, nous passons en revue chacune des fonctionnalités du LMCE. Pour chacune d'elles, l'objectif est de mettre en relation le(s) élément(s) d'interface et le(s)

modèle(s) conceptuel(s) impliqué(s). Ces modèles conceptuels sont ici assimilés à des bases de données, de façon à mettre en lumière les échanges de données entre le système et ces modèles. Les autres types de données sont également traités comme des bases de données. L'objectif est d'aboutir à une architecture générale du LMCE.

Ces fonctionnalités peuvent être représentées sous forme d'activités. Les éléments d'interface et les bases autres que celles relatives aux modèles conceptuels sont alors les ressources de ces activités. Les modèles conceptuels fournissent les informations utilisées comme flux de ces activités. Des diagrammes d'activité détaillent les flux impliqués dans ces différentes fonctionnalités. L'énumération suivante présente ces fonctionnalités.

Créer ou modifier les profils des utilisateurs : ces fonctionnalités mettent en relation le formulaire de saisie des profils et la base des profils utilisateurs. Dans les deux cas, la base est interrogée, soit pour vérifier que le profil créé ne concerne pas un utilisateur déjà existant, soit pour obtenir les données du profil à modifier. Une fois les formulaires validés, la base des profils est mise à jour.

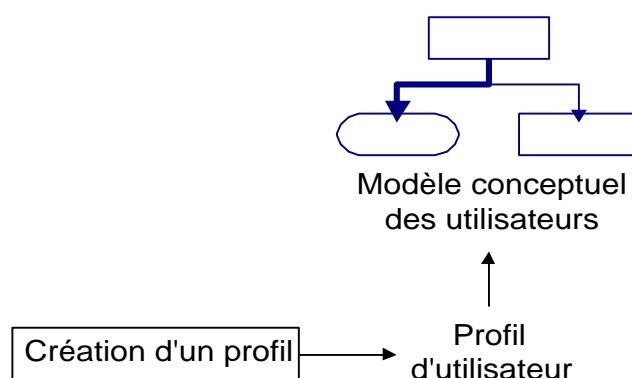


Figure 37. Fonctionnalité de création d'un profil.

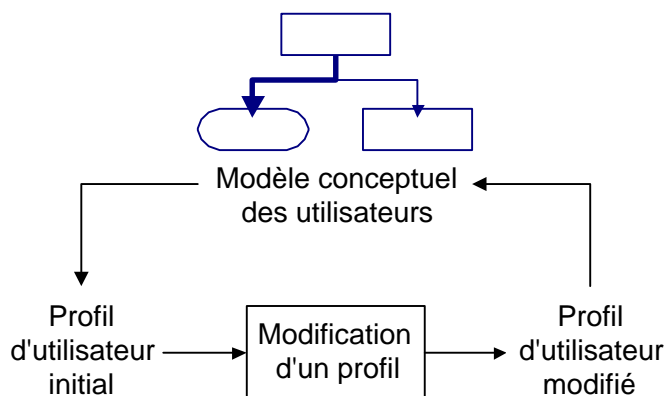


Figure 38. Fonctionnalité de modification d'un profil existant.

Autoriser ou refuser la connexion d'un utilisateur au LMCE : cette fonctionnalité, réalisée par le formulaire de connexion qui interroge la base des profils.

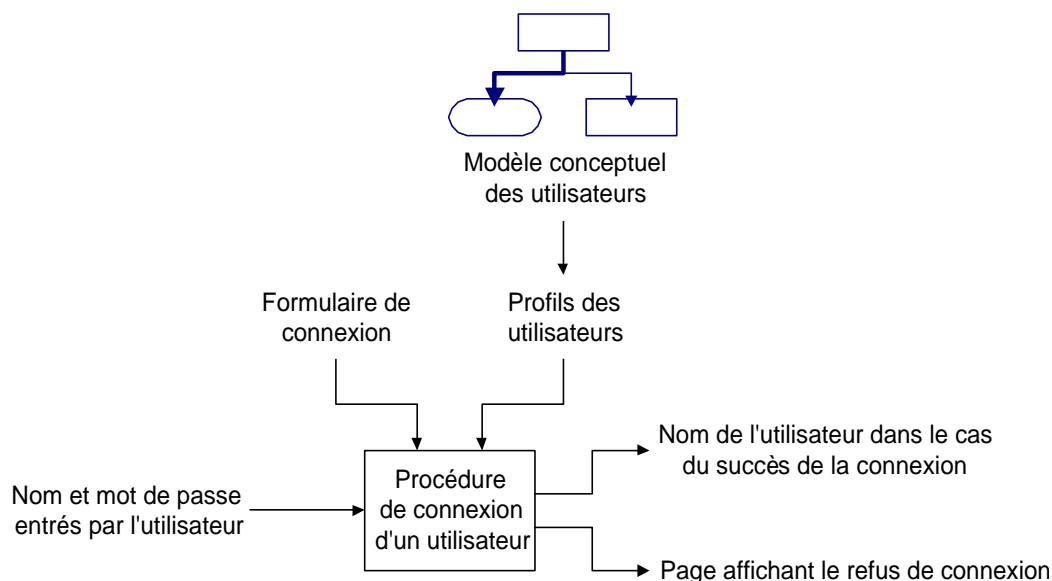


Figure 39. Acceptation ou refus de connexion d'un utilisateur au système.

Sélectionner le type d'activité générique d'utilisation du LMCE correspondant au profil de l'utilisateur : cette fonctionnalité, réalisée une fois la connexion validée, nécessite l'interrogation de la base des profils.

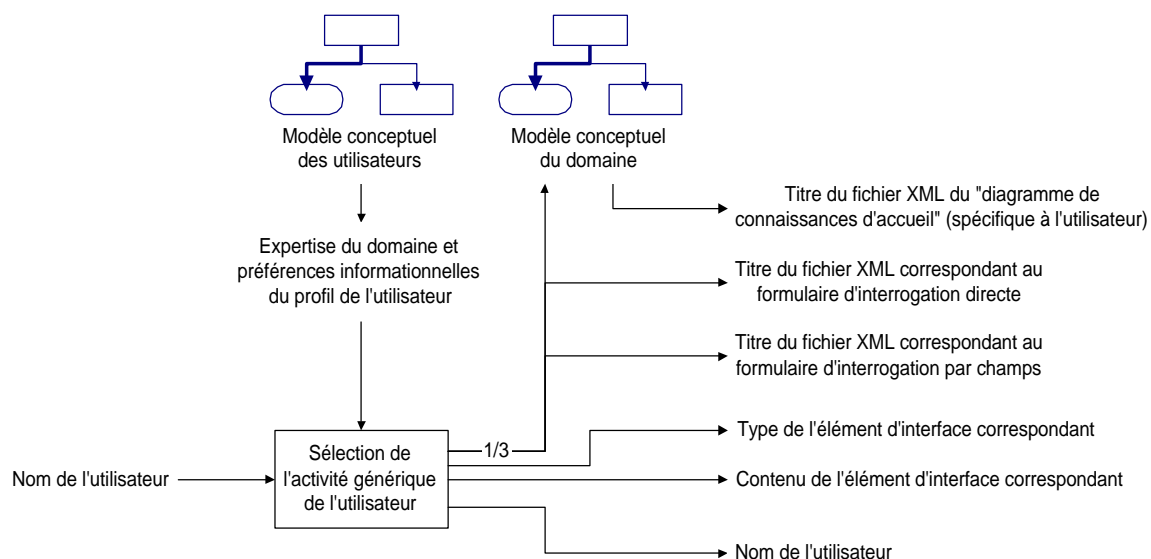


Figure 40. Sélection de l'activité générique de l'utilisateur du LMCE. Les nom, type et contenu de l'élément HTML à afficher pour cette activité sont transmis à la fonction d'affichage. Le nom de l'utilisateur est propagé car il est nécessaire aux fonctionnalités interrogeant la base des profils.

Afficher les éléments de l'interface correspondant à une fonctionnalité : cette fonction générique est appelée à chaque changement des informations présentées à l'écran. Elle permet d'assurer l'homogénéité des différentes « pages écran ». Elle interroge la base des profils utilisateurs et les bases nécessaires à l'affichage des différents composants de l'interface.

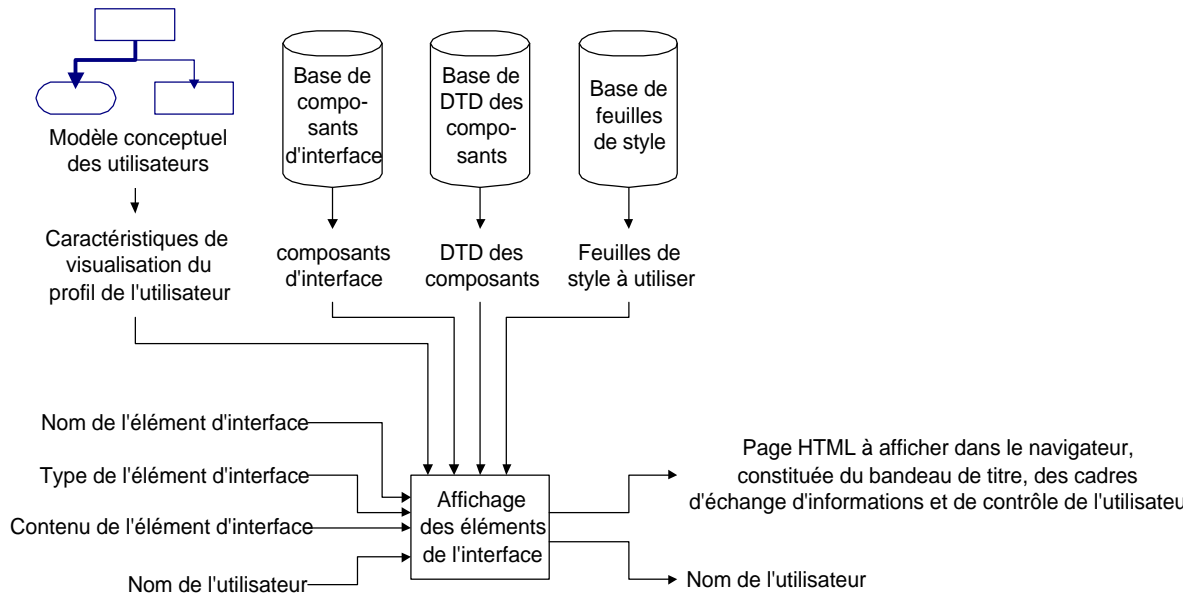


Figure 41 Fonctionnalité générique d'affichage du LMCE. Cette fonctionnalité permet l'affichage des diagrammes de connaissances, mais également des différents formulaires, des réponses du moteur de recherche et des documents.

Permettre la navigation hypermédia dans ces diagrammes : cette fonctionnalité implique le cadre de présentation des diagrammes de connaissances, les multiplexeurs de liens et les cartes conceptuelles. La base de diagrammes est interrogée.

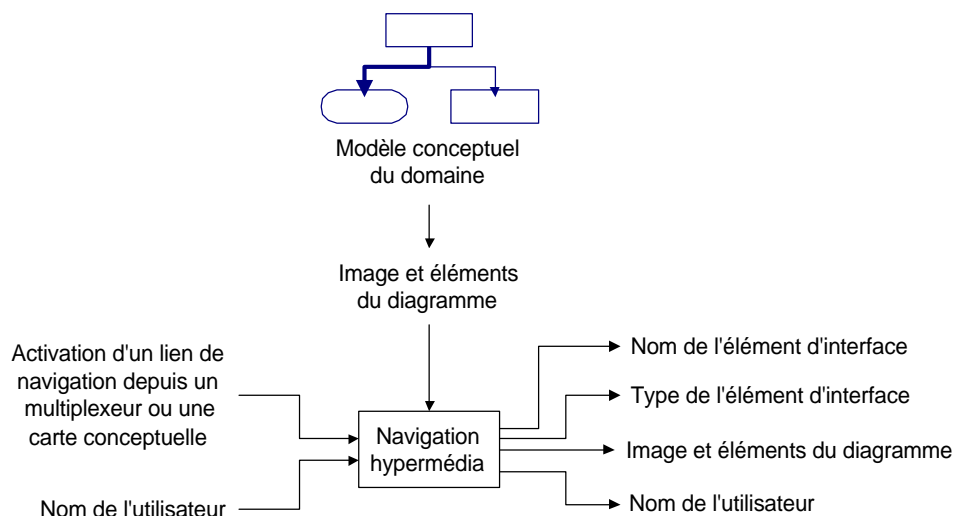


Figure 42. Fonctionnalité générique de navigation hypermédia avec le LMCE. Elle est déclenchée par l'évènement "onclick" sur un lien d'un diagramme.

Permettre la sélection d'éléments des diagrammes : cette fonctionnalité est réalisée par les multiplexeurs de liens et le caddie à concepts. Elle ne fait pas appel à un modèle conceptuel du LMCE, car elle utilise les éléments du cadre HTML généré par la fonctionnalité précédente.

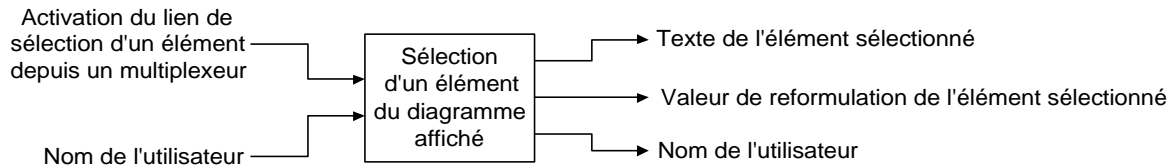


Figure 43. Fonctionnalité de sélection d'un élément d'un diagramme pour l'accès à l'information cumulatif.

Permettre l'articulation des textes des éléments sélectionnés : cette fonctionnalité est réalisée en interne par le caddie à concept. Elle n'est pas détaillée dans un diagramme, car elle ne met pas en cause de flux particuliers.

Formaliser des requêtes interrogatives complexes à partir des éléments sélectionnés : cette fonctionnalité est réalisée en interne par le caddie à concepts.

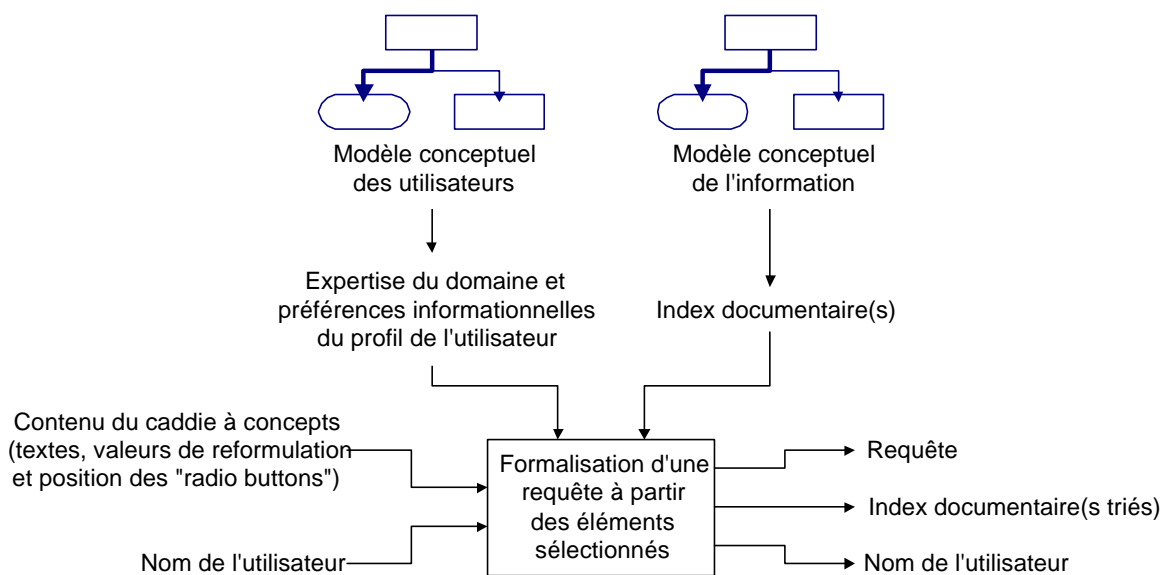


Figure 44. Formalisation d'une requête selon le langage d'interrogation du moteur de recherche utilisé à partir du contenu du caddie à concepts.

Permettre la formulation de requêtes interrogatives directes ou par champs : cette fonctionnalité est réalisée en interne par les formulaires correspondants. Comme pour l'articulation des éléments de la requête pour l'interrogation cumulative, cette fonctionnalité se limite à la présentation des éléments HTML adéquats et n'est pas détaillée ici.

Interroger le fonds documentaire de l'organisation en fonction de l'indexation conceptuelle de ce fonds : cette fonctionnalité est la plus complexe du LMCE. Elle consiste à

utiliser la requête du formulaire d'interrogation courant pour interroger la base des profils des utilisateurs et le modèle conceptuel du système d'information (en sélectionnant les index correspondants à la requête). Cette fonctionnalité est réalisée par le moteur de règles.

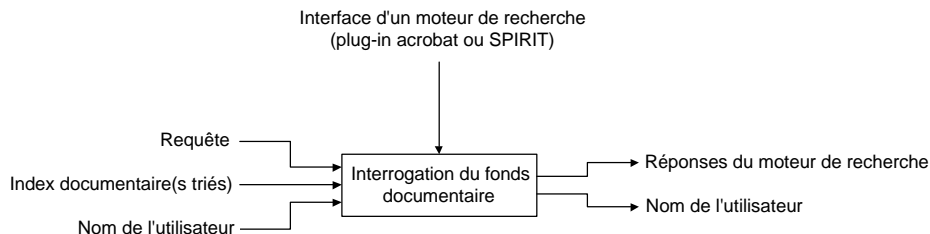


Figure 45. Envoi de la requête au moteur de recherche pour l'interrogation des index sélectionnés.

Présenter les résultats de l'interrogation du fonds documentaire : cette fonctionnalité est pour l'instant réalisée en interne par le plug-in Acrobat Search® d'Adobe et n'implique aucun élément de l'interface. Dans l'hypothèse où le fonds est indexé et interrogé par SPIRIT, cette fonctionnalité est visualisée dans un cadre de l'interface du LMCE. Dans ces deux cas, elle ne fait appel à aucun modèle conceptuel de cet outil.

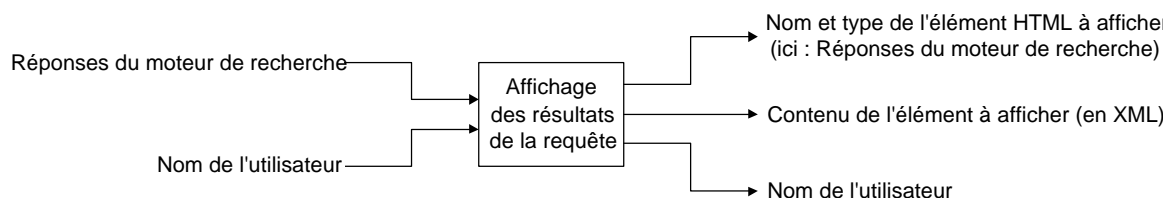


Figure 46. Fonctionnalité de prétraitement des résultats de la recherche renvoyés par l'interface du moteur de recherche pour la fonction générique d'affichage.

Afficher un document : cette fonctionnalité fait directement appel à la base documentaire externe de l'organisation et non au modèle conceptuel du système d'information, qui est lui constitué des index utilisés pour l'interrogation.

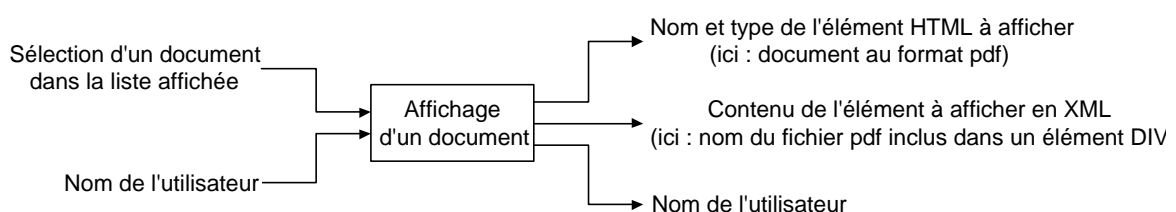


Figure 47. Même fonctionnalité que précédemment pour l'affichage des documents.

Gérer les paramètres spécifiques de visualisation des utilisateurs (à terme) : cette fonctionnalité vise à permettre la visualisation des différents éléments d'interface en fonction des préférences spécifiques de chacun. Elle n'est pas réalisée ici. Elle ferait notamment appel à un éditeur et à une base de feuilles de styles. Actuellement, les feuilles de style sont supposées communes à tous les utilisateurs. Nous avons tout de même décrit cette fonctionnalité dans le diagramme d'activités ci-dessous.

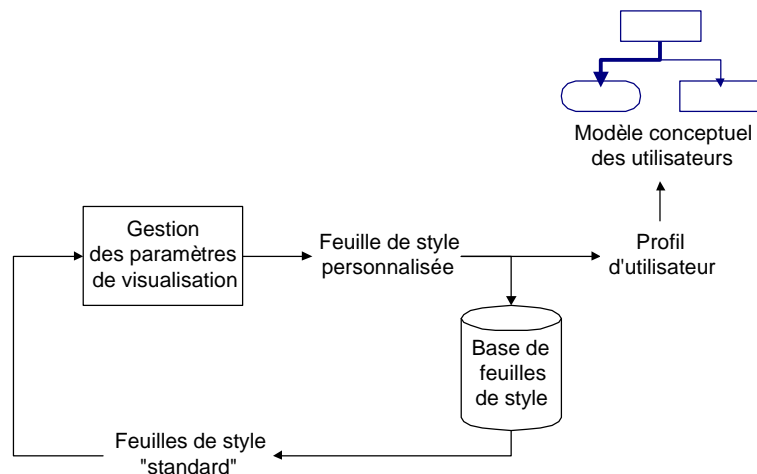


Figure 48. Fonctionnalité de mise à jour de la base de feuilles de styles pour la gestion des paramètres de visualisation personnalisés.

Chacun des points ci-dessus détaille l'une des fonctionnalités du LMCE. Le diagramme suivant présente la séquence de ces fonctionnalités dans le fonctionnement global du LMCE. L'activité correspondant à la stratégie d'accès à l'information cumulative est détaillée dans un second diagramme. Pour éviter de les surcharger, nous avons omis les flux échangés entre les différentes activités. Seuls les états de début et de fin des séquences d'activités sont mentionnés. La séquence principale du fonctionnement du LMCE comporte un début mais pas de fin. En effet, le système boucle sur la navigation hypermédia ou l'accès à l'information. La fin de la session est déterminée par la fermeture du navigateur.

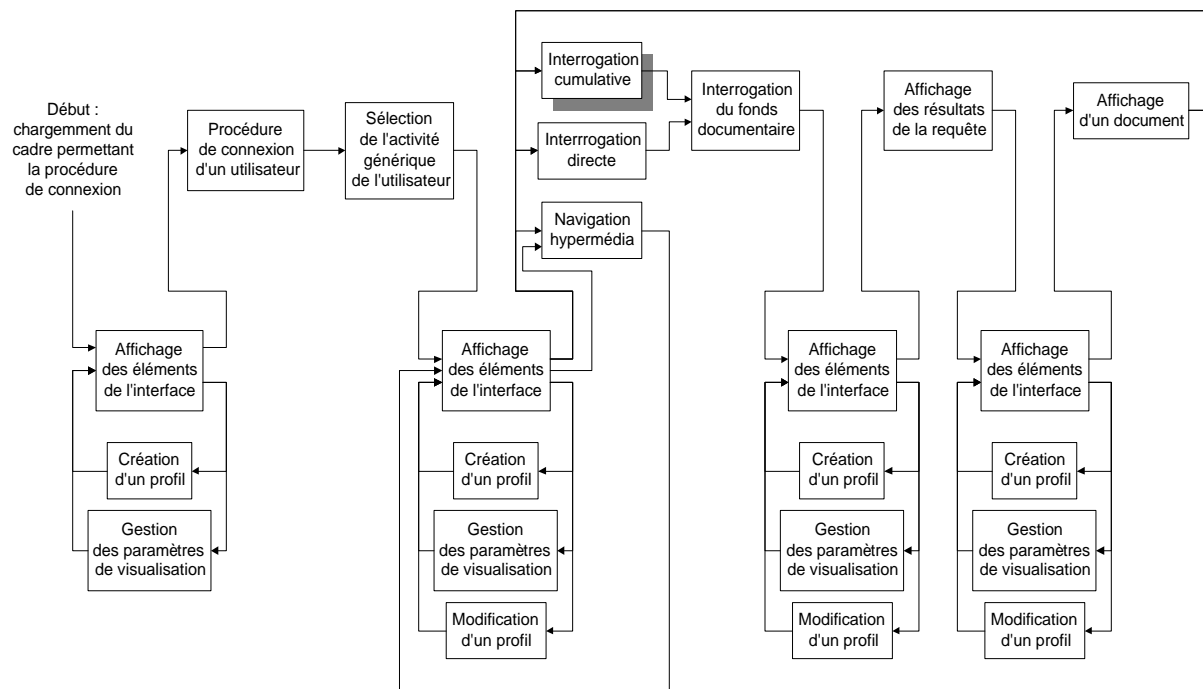


Figure 49. Séquence des activités dans le fonctionnement général du LMCE.

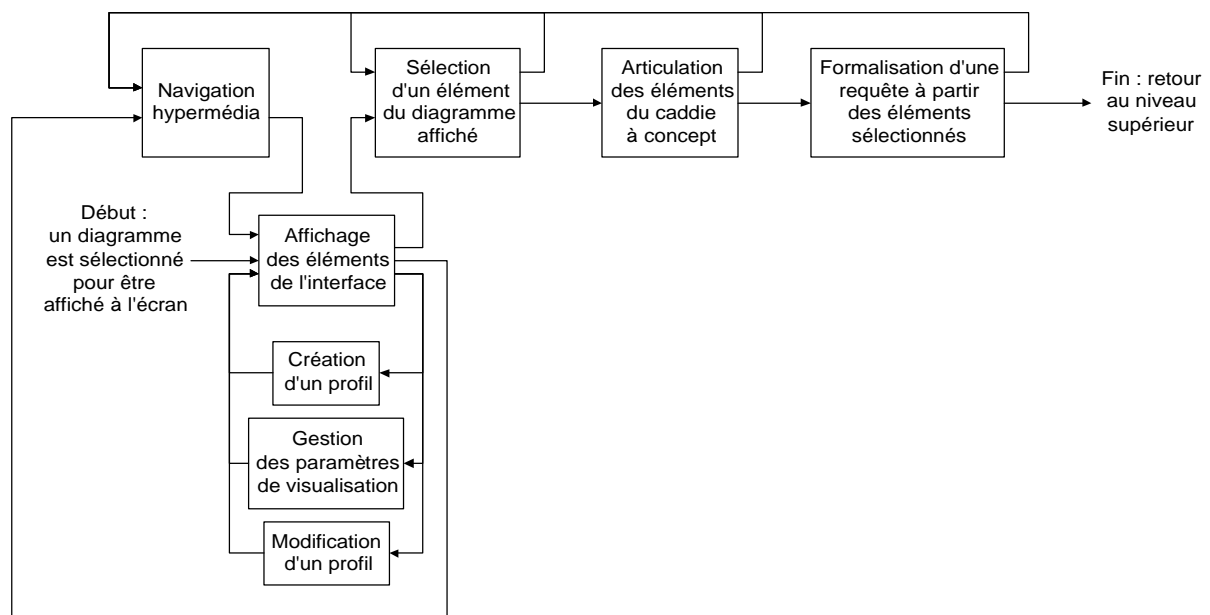


Figure 50. Détails des activités liées à la stratégie d'accès à l'information cumulative.

Les différentes fonctionnalités décrites ci-dessus sont réalisées par des scripts en langage JSCRIPT incorporés aux différents éléments de l'interface. Le tableau ci-dessous indique à quel élément de l'interface correspond chaque fonctionnalité.

Fonctionnalité	Emplacement
Connexion	Formulaire de connexion
Création/modification d'un profil	Formulaire de création/modification d'un profil
Sélection de l'activité générique de l'utilisateur	Formulaire de connexion
Affichage des éléments de l'interface	Page principale
Navigation	Multiplexeur de liens
Affichage du multiplexeur de liens	Diagramme de connaissances affiché
Sélection d'un élément d'un diagramme	Multiplexeur de liens
Formalisation d'une requête à partir de l'expression du besoin d'information	Caddie à concepts
Envoi de la requête	Caddie à concepts
Affichage des résultats de la requête	Caddie à concepts
Affichage d'un document	Cadre d'affichage des résultats de la requête
Mise à jour des préférences de visualisation	Cadre de mise à jour des préférences de visualisation (non encore réalisé)

Tableau 5. Emplacements des différentes fonctionnalités de l'interface. Les fonctionnalités d'affichage, qui ne concernent que l'appel de la fonction générique d'affichage, sont réalisées dans l'élément d'interface précédent. La fonctionnalité de formalisation des requêtes pour l'interrogation cumulative fait appel au système de règles. Celle d'envoi de la requête utilise le contrôle ActiveX d'appel du plug-in Acrobat.

4.2.5.2. Architecture du LMCE

À partir des fonctionnalités décrites au paragraphe précédent, l'architecture générique du LMCE réalisant ces spécifications est simple. Il s'agit alors d'un outil composé des trois modèles conceptuels décrits au début de ce chapitre et de bases de données permettant de stocker les contenus (objets visualisables et scripts) des composants de l'interface, les DTD et les feuilles de style associées.

Cette simplicité est rendue possible par l'approche DVP que nous avons employée tout au long de ce processus de conception. Cette approche nous a surtout permis de mettre au point cette architecture modulaire et d'incorporer dans des scripts les fonctionnalités du système. Le schéma suivant positionne cette architecture dans le fonctionnement général du système.

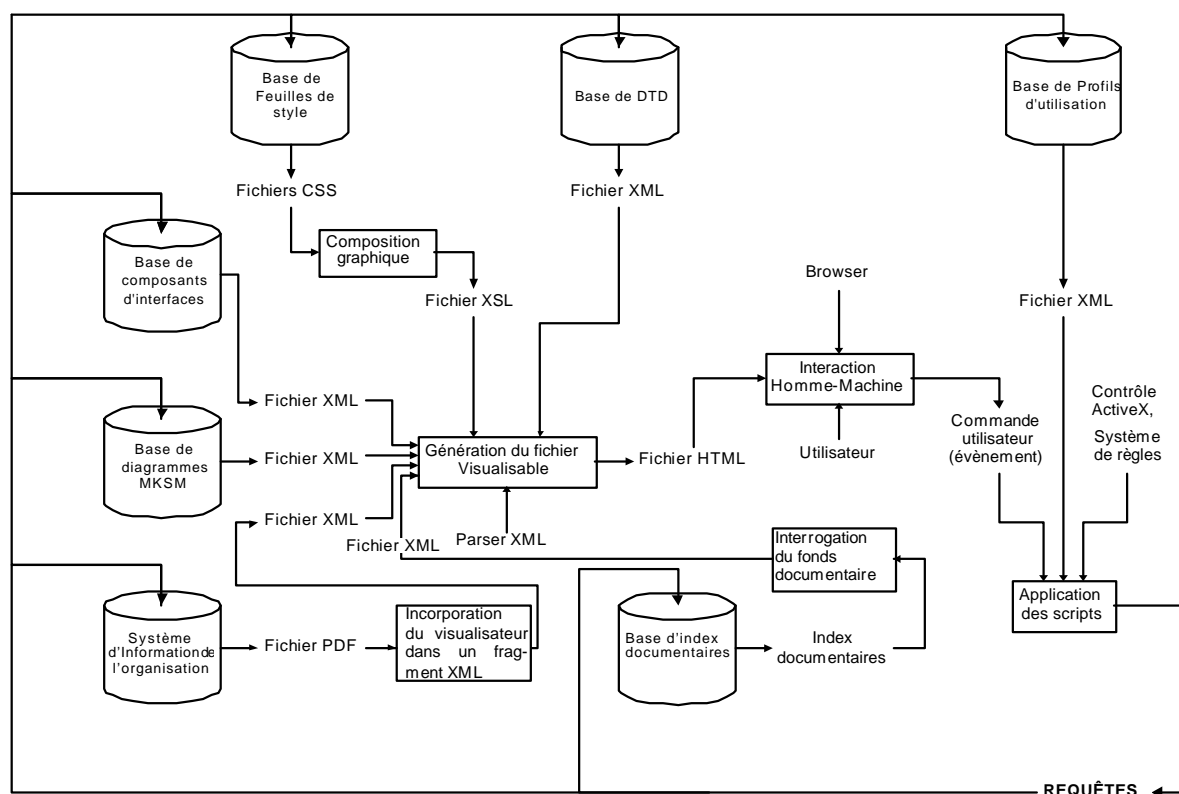


Figure 51. Architecture et fonctionnement du LMCE.

4.2.5.3. Éléments complémentaires des spécifications techniques

Dans ce paragraphe, nous apportons des éléments d'information complémentaires, concernant les éléments de cette architecture. Les aspects techniques de réalisation des différentes bases de données et le moteur de règles sont décrits.

4.2.5.3.1. Les bases de données

Dans la figure ci-dessus, toutes les sources d'informations (structurées ou non structurées) sont représentées sous formes de bases de données. Cela permet d'exprimer le fait que pour chacun de ces types d'information, il existe plusieurs éléments disponibles et qu'il est possible d'y ajouter des éléments. Cependant, dans ce prototype, peu d'informations sont réellement stockées dans des bases de données. En effet, XML permet par définition de réaliser sur les données les opérations effectuées par les bases de données. Le DOM HTML (DOM Core et interfaces pour le langage HTML) permet de plus de réaliser ces opérations avec des coûts de développement bien inférieurs. Nous tirons ici avantage du fait que le parser MSXML est inclus dans le navigateur Microsoft Internet Explorer® 5.

C'est pourquoi nous avons choisi, pour ce prototype, d'accéder directement aux données contenues dans des documents XML stockés sous forme de fichiers. Les requêtes et les mises

à jour des bases sont alors de simples accès aux fichiers correspondants. La contrainte inhérente à ce choix concerne la taille des structures de données, qui doit rester « raisonnable »¹²³. Compte tenu du fait que la plupart de nos bases contiennent des documents XML au format texte dont les fragments sont des éléments HTML, cette contrainte n'est pas significative ici. Dans ce qui suit, nous passons en revue ces « bases » et indiquons, pour chacune d'elles, sa fonction par rapport à la finalité du LMCE. Nous avons classé ces données en deux types : informations et méta-données. Les informations sont les données effectivement visualisables, tandis que les méta-données sont nécessaires au fonctionnement du LMCE mais ne sont pas signifiantes pour l'utilisateur.

4.2.5.3.1.1. Informations

La base de données en criticité représente le système d'information de l'organisation. Comme nous l'avons déjà souligné, la recherche est multi-bases et d'autres formats de fichiers sont donc envisageables. La visualisation se fait par l'intermédiaire d'un outil spécifique dont la liste est gérée par le navigateur. Le fonds pdf correspond à l'exemple traité dans l'appliquatif de cette thèse et est plus contraignant que la plupart des formats de fichiers courants. La base de données en criticité contient donc l'ensemble des informations auxquelles les utilisateurs vont avoir besoin d'accéder via le LMCE.

La base de diagrammes MKSM représente le modèle du domaine présenté au début de ce chapitre. Elle est constituée d'un fichier XML par diagramme et d'un autre fichier XML indiquant les titres, les formalismes et les noms de ces fichiers, répertoriés par sous-domaines. Ce dernier fichier est utilisé pour accéder aux fichiers des diagrammes lors de la navigation (ce procédé évite de conserver une liste de tous les noms de fichiers dans un script) et pour indiquer les références du diagramme visualisé dans le bandeau de titre.

4.2.5.3.1.2. Méta-données

La base de composants d'interface est composée d'un fichier XML contenant sous forme de fragments les différents formulaires décrits précédemment et la structure générique du bandeau de titre et du multiplexeur de liens. Pour des raisons de facilité de maintenance (voir plus loin), les cartes conceptuelles sont stockées séparément. C'est également dans cette base que se trouveraient les structures des cadres d'édition des préférences graphiques, de visualisation des réponses du moteur de recherche W3 et de visualisation des documents. Les

¹²³ Nous définissons ici ce terme comme permettant la manipulation et l'édition de ces fichiers avec un éditeur de

scripts inclus dans ces éléments sont également inclus dans ces fragments. Ils sont au format XML/Javascript pour MS IE4.0.

La base de feuilles de style permet de spécifier et de différencier l’affichage de chaque élément de l’interface. Des feuilles de style CSS génériques permettent d’homogénéiser l’interface globale du LMCE. Les fragments XML tels que les éléments d’interface sont traités par des feuilles de style XSL. Les plus complexes (c’est-à-dire ceux dont le fragment XML ne contient que la structure) font appel à des transformations avec XSLT.

La base de DTD n’est actuellement pas réalisée. Seules les références aux entités utilisées dans les fichiers XML sont pour l’instant présentes. Cette base est destinée à permettre au parser utilisé de valider les documents XML des autres bases. Compte tenu des limitations des DTD (notamment au niveau du typage des données), nous réfléchissons actuellement à l’utilisation de schémas XML pour la validation de ces bases. Ces schémas utilisent la même structure que les documents XML eux-mêmes et sont donc plus puissants et plus dynamiques que les DTD. De plus, l’utilisation de schémas permet une structuration par objet du même type que celle que nous avons adoptée pour les « bases » XML.

La base de profils utilisateurs contient un document XML dans lequel chaque utilisateur est stocké sous forme de fragment. La structure de ces fragments est celle indiquée dans le paragraphe de définition du modèle conceptuel de l’utilisateur.

La base d’index documentaires n’est en fait, comme nous l’avons déjà signalé, qu’une arborescence de répertoires contenant les différents index permettant l’interrogation du fonds. Un document XML à la racine de ce répertoire permet d’obtenir les chemins absolus de ces index à partir de leurs noms. Ces chemins sont récupérés par le script du formulaire d’interrogation et passés par arguments à l’appel du contrôle ActiveX d’interrogation.

4.2.5.3.2. *Les contenus dynamiques*

Nous avons vu dans le paragraphe du même nom que l’ensemble du système de règles nécessaire au fonctionnement du LMCE est en fait constitué de scripts disséminés dans les différents éléments d’interface. De la même manière, le moteur de règles qui gère l’interrogation du fonds documentaire est inclus dans le fragment XML de la base de composants d’interface qui concerne le caddie à concepts.

Le contrôle ActiveX est incorporé dans les formulaires d'interrogation sous la forme des boutons « Envoyer la requête ». Celui-ci valide ces index dans la liste des index d'Acrobat Search, soumet la requête et les retire de la liste.

4.2.6. Cycle de vie du LMCE

Le prototype tel qu'il est spécifié ci-dessus permet la navigation dans les diagrammes de connaissances et l'accès à l'information grâce à la modélisation réalisée dans le projet-pilote. Nous avons vu que le LMCE est conçu pour pouvoir être réutilisé dans d'autres domaines et à partir des modélisations effectuées dans d'autres organisations. Dans cette optique, nous étudions ici deux phases du cycle de vie de cet outil : sa génération et sa maintenance.

4.2.6.1. Génération du LMCE

L'approche DVP qui a conduit à l'architecture modulaire de cet outil permet sa réutilisabilité en séparant les données des structures de données. Toutefois, toutes les données ne doivent pas nécessairement être générées pour chaque nouvelle instance du LMCE. Il s'agit dans ce paragraphe de rappeler quels sont les éléments du LMCE qui sont réutilisables et de prévoir une méthodologie spécifique de génération des composants permettant d'aboutir, le plus automatiquement possible, à une instance de LMCE.

Comme nous l'avons indiqué précédemment, seuls les modèles conceptuels du domaine et de l'information sont spécifiques à l'organisation modélisée. Nous admettons pour l'instant que la structure du modèle des utilisateurs est conservée, et que par conséquent, le système de règles réalisant notamment le filtrage de l'information n'a pas à être modifié. Chaque instance d'un profil est créée à la première connexion d'un utilisateur au système. Cette assumption est fondée sur la relative généricité des profils présentés ici. Le cas d'une modification de la structure des profils d'utilisateurs nécessitant la réécriture spécifique de chacune des règles n'est pas traité ici. Il est abordé à un niveau méthodologique dans le paragraphe concernant les améliorations possibles du système au chapitre 5.

Les parties précédentes de ce chapitre indiquent la méthodologie de mise au point des différents modèles. Pour chaque projet, nous avons vu que la rédaction des diagrammes de connaissances et que l'indexation des sources documentaires pouvaient être effectuées en parallèle. Ces modélisations permettent d'obtenir **les bases de diagrammes MKSM et des index documentaires** indiquées dans le paragraphe précédent.

La base de DTD doit être modifiée pour définir les références aux sous-domaines modélisés. L'utilisation de schémas XML plutôt que celle de DTD présente ici l'avantage d'éviter la modification de cette base. En effet, les schémas XML sont des documents XML capables d'identifier dynamiquement les différents sous-domaines référencés dans les diagrammes et de définir les références nécessaires en conséquence (cf. [Hunter & al. 01], chap. 7, pour plus de précisions sur l'utilisation des schémas XML).

La base de composants d'interface contient les cartes conceptuelles qui doivent être générées spécifiquement en fonction de la modélisation réalisée. La méthode de réalisation de ces cartes conceptuelles n'est pas traitée ici. Cependant, les ateliers de conception réalisés jusqu'ici autour de la méthode MKSM ont montré qu'il est possible d'automatiser le processus de génération de ces cartes.

La base des profils est vide à l'origine et n'a pas à être générée spécifiquement pour chaque projet.

La base des feuilles de style (par défaut) est générique.

La base de données documentaire de l'organisation préexiste au LMCE et n'est pas concernée par la génération de cet outil.

Le contrôle ActiveX d'interrogation du fonds documentaire et par extension les API des autres systèmes de recherche d'information utilisés font par définition le lien entre le LMCE et les moteurs de recherche. Ils sont donc indépendants du contenu indexé et chaque API existante peut être réutilisée dans un projet ultérieur. Il est par contre nécessaire de mettre au point les API des systèmes non utilisés dans les projets précédents.

Le moteur de règles est générique et n'a pas à être modifié pour chaque projet. Moyennant les assumptions du début de ce paragraphe, les règles sont également indépendantes de la modélisation.

L'automatisation de la production de LMCE consiste à proposer des outils de réalisation des différents éléments à modifier identifiés ci-dessus. Certains de ces outils existent, et sont utilisés pour la conception de LMCE. C'est par exemple le cas du logiciel Visio®, que nous avons utilisé pour la rédaction des diagrammes. Ce peut également être le cas pour l'atelier de génération de livres de connaissances MNESIK® de C-Log, que nous n'avons pas eu le temps de tester. Cet atelier, comme l'atelier MKSM précédent, permet d'ailleurs également de mettre au point les cartes conceptuelles spécifiques à MKSM. De même, les outils d'indexation du fonds documentaire sont fournis par l'éditeur du moteur de recherche utilisé

(par exemple, Acrobat Catalog® pour Acrobat Search®). Les autres éléments du LMCE à mettre à jour sont constitués de fichiers XML.

Nous proposons donc une « chaîne de conception » de LMCE en deux étapes. La première est la mise au point préalable des éléments à l'aide des outils cités ci-dessus. La seconde fait appel à une application composée d'une séquence de formulaires permettant de définir un à un les différents fragments des documents XML à générer. Cette application s'appuie sur les structures génériques des documents XML proposés dans ce manuscrit. Elle n'est pas spécifiée ici.

4.2.6.2. Maintenance du LMCE

Une fois le LMCE généré tel que cela est préconisé au paragraphe précédent, cet outil est destiné à être livré et installé en intranet dans l'organisation modélisée. Nous ne présumons pas de la durée de vie de ce système, mais il est fort probable qu'il doive être remis à jour pendant sa durée d'utilisation. Plusieurs types de situations peuvent nécessiter la maintenance du LMCE. Nous avons identifié deux types de maintenance pour ce système.

Maintenance synchrone : il s'agit ici d'une maintenance effectuée dans le cadre du fonctionnement « normal » de l'organisation. Ce fonctionnement induit une production et une consommation d'information. Les nouveaux documents doivent être ajoutés aux bases et les documents éventuellement jugés obsolètes (comme dans le cas du remplacement d'une ancienne version d'un document par sa nouvelle version) doivent en être exclus. Ce type de maintenance concerne donc la mise à jour de l'indexation conceptuelle. Les seuls composants du LMCE affectés sont les index documentaires. Il s'agit alors, pour chaque document ajouté, de définir les critères d'indexation auxquels il correspond et de mettre à jour les index correspondants. De même, les documents obsolètes doivent être supprimés de tous les index dans lesquels ils figurent¹²⁴. Acrobat Catalog® et SPIRIT® 1.6 permettent l'ajout et la suppression d'un document dans une base. Cette maintenance synchrone peut donc être effectuée en interne par le personnel de l'organisation modélisée. Toutefois, compte-tenu du nombre d'index, cette mise à jour des bases peut être fastidieuse. C'est pourquoi il paraît intéressant de l'automatiser, par exemple en réalisant une application permettant, à partir du nom d'un fichier et de ses descripteurs, de réaliser la mise à jour de tous les index correspondants. Une telle application n'est pas proposée ici.

¹²⁴ Une liste inverse du contenu des bases pourra être tenue à jour. Il s'agit d'inverser la liste des documents contenus dans chaque base, pour obtenir celle des bases contenant chaque document.

Maintenance diachronique : des changements plus profonds peuvent affecter l'utilité du LMCE pour l'organisation. C'est par exemple le cas d'une modification de l'activité de l'organisation, due à l'arrivée de nouvelles méthodes ou à l'acquisition de nouveaux outils. Cela peut également se produire dans le cas d'une restructuration interne ou d'une modification du Plan d'Assurance Qualité. Dans ce cas, le modèle du domaine du LMCE doit être remis à jour. En fonction de l'importance de ces changements, deux types d'opérations de maintenance sont à distinguer.

Une mise à jour limitée de certains diagrammes de connaissances concernant les aspects de la modélisation à actualiser peut intervenir. Il s'agit alors de « régénérer » la collection de diagrammes, les DTD, les valeurs de reformulation des éléments des diagrammes et les cartes conceptuelles (que nous avons séparés des autres composants d'interface). Ce travail doit être fait par un chargé de projet en gestion des connaissances, mais se limite à une mise à jour de la modélisation.

Une réactualisation complète de l'outil peut intervenir en cas de changements plus profonds dans l'activité de l'organisation. Nous considérons que de tels changements sont avérés si la modification de l'activité influe sur l'indexation conceptuelle du système d'information de l'organisation. Dans ce cas, les critères d'indexation sont à redéfinir avec les « nouveaux experts¹²⁵ » du domaine et tout le fonds documentaire doit être réindexé en fonction de ces nouveaux critères. Cette redéfinition du modèle conceptuel du système d'information s'accompagne de la refonte complète du modèle du domaine, selon la méthode indiquée au paragraphe précédent. Au final, ce type de maintenance revient à définir un nouveau LMCE en s'« aidant » de la modélisation réalisée dans le projet précédent.

Si la maintenance synchronique du LMCE ne pose pas de problème conceptuel particulier, nous n'avons pas de réponse simple à proposer pour la maintenance du LMCE dans le cas de modifications profondes de l'activité de l'organisation. L'intérêt même de ce type de système peut être remis en cause dans le cas du changement d'activité de l'organisation.

¹²⁵ Cette locution contradictoire désigne les individus qui font désormais autorité dans le domaine. Le domaine ayant changé, ces individus sont qualifiés de nouveaux pour les différencier des experts de l'ancien domaine. Cependant, l'expérience des projets prouve que l'expertise s'acquiert avec le temps. On ne peut donc a priori pas être expert dans un domaine nouveau. Dans ces conditions, nous nous interrogeons sur le sens à donner à cette locution. Il est donc peut-être possible de ne pas trouver d'expert capable de renseigner le chargé de projet. Dans ce cas, il n'y a pas de savoir-faire à modéliser et la question de la pertinence de la démarche du LCE se pose.

4.3. Application au projet-pilote

Les parties précédentes nous ont permis d'exposer la structure conceptuelle du livre multi-utilisateurs de connaissances électronique, la démarche ergonomique de conception de cet outil, ses spécifications et son cycle de vie. Cette étude s'appuie sur le projet de gestion des connaissances mené au SEC de l'IPSN dans le domaine de la maîtrise du risque de criticité hors réacteur nucléaire. L'objectif de cette partie est de présenter l'état de réalisation de la modélisation et du prototype à la date de la fin de la thèse au CEA. Pour cela, nous indiquons l'état d'achèvement des modèles conceptuels et des fonctionnalités de l'outil à cette date.

4.3.1. Les modèles conceptuels

Nous décrivons dans ce paragraphe les trois modèles conceptuels sur lesquels s'appuie le prototype. Ces modèles ont été réalisés en grande partie pendant la phase de modélisation au SEC et immédiatement après cette phase, qui s'est déroulée au début de la thèse. À la date de rédaction de ces modèles, l'état d'avancement de la réflexion sur le LMCE n'était pas celui présenté dans le reste de ce chapitre. Depuis, les spécifications du LMCE ont été mises au point. C'est pourquoi ces modèles peuvent différer des modèles finaux présentés précédemment dans ce chapitre.

Par ailleurs, il a été envisagé au début de la thèse d'utiliser le LMCE pour d'autres activités que l'accès à l'information. Ces activités sont indiquées dans le modèle des utilisateurs. Nous avons finalement renoncé à cette idée. Les raisons qui nous ont conduit à ce changement sont exposées au paragraphe 4.3.3. L'objectif était d'inclure ces applications dans un cadre du LMCE, de la même manière que les applications permettant la visualisation des informations accédées. Si cette fonctionnalité n'a pas été retenue, le raisonnement correspondant permet toutefois de repositionner le LCE comme un élément du cycle de vie de l'information. Il se justifie donc par la mise en lumière des objectifs globaux des différents types d'utilisateurs et est indiqué ici. Pour cette raison, les outils du système d'information sont également modélisés dans le modèle du système d'information.

4.3.1.1. Le modèle du domaine

L'état de la modélisation des connaissances au SEC est indiqué au chapitre précédent : Elle comporte environ une centaine de diagrammes. Deux sous-domaines sont suffisamment modélisés pour permettre la navigation et l'accès à l'information à partir des diagrammes.

4.3.1.2. Le modèle des utilisateurs

Lors de la phase de capitalisation des connaissances du projet, nous avons identifié six sous-domaines de la sûreté-criticité. Deux d'entre eux sont des sous-domaines d'ingénierie et deux autres des domaines de recherche. Les caractéristiques des utilisateurs experts étant les mêmes dans chacune de ces deux classes de sous-domaines, ils ne sont pas différenciés. Les deux autres sous-domaines ne concernent qu'un seul individu et n'ont pas donné lieu à la modélisation d'un type d'utilisateur particulier.

La liste des sous-domaines et les trois niveaux d'expertise définissent les différents types d'utilisateurs du prototype en sûreté-criticité. En parallèle de la capitalisation des connaissances, nous avons recueilli des informations sur les façons de travailler (buts du travail, types d'informations recherchées, outils actuels d'accès à l'information) des types d'utilisateurs experts modélisés.

Ces informations nous ont permis d'identifier les caractéristiques de ces types d'utilisateurs. À partir de celles-ci, nous avons défini les profils d'utilisation génériques par défaut présentés plus haut. Les caractéristiques des utilisateurs correspondants sont présentées ci-dessous.

4.3.1.2.1. *Expert Ingénieur*

4.3.1.2.1.1. **Besoins**

Un ingénieur a en général besoin d'accéder le plus vite possible à une donnée particulière, comme la valeur minimale critique correspondant à la configuration qu'il étudie et de l'utiliser pour un calcul de criticité via l'interface du code de calcul standard utilisé pour les études. Il peut ensuite avoir besoin d'inclure le résultat obtenu dans une note rédigée selon un formalisme prédéfini et par l'intermédiaire d'un logiciel de traitement de texte. L'enchaînement des tâches accomplies par un ingénieur et des informations nécessaires est relativement bien décrit par les diagrammes d'activité MKSM. Pour être utile à l'ingénieur, notre outil doit intégrer et relier les différents stades de production documentaire (recherche d'information, calcul, rédaction) pour lui permettre de gagner du temps.

4.3.1.2.1.2. **Caractéristiques particulières**

- bonne connaissance des paramètres et des mécanismes utilisés dans les calculs de criticité, ainsi que des outils à sa disposition
- connaissance moyenne des phénomènes physiques (neutroniques) liés à la criticité,

- connaissance de la terminologie technique et des références les plus utilisées,
- pragmatisme (i.e. « time pressure »).

4.3.1.2.1.3. Tâches

Navigation : la navigation se fait essentiellement à partir des modèles dynamiques (activité, tâches), car ce sont ces modèles qui décrivent le processus de production de ces utilisateurs. En suivant séquentiellement ces modèles dynamiques, l'ingénieur arrive au terme d'une production documentaire prédéterminée. En fonction du degré de finesse atteint dans la décomposition de ces activités, il accède à des modèles statiques qui identifient les données de sa problématique (objets, concepts). Le mode de navigation à travers les liens statiques privilégié pour ce type d'utilisateur est donc une navigation « séquentielle ».

Recherche d'information : la recherche d'information se fait en fonction des deux types de critères (i.e. de forme et de fond) identifiés dans le modèle des sources d'informations du SEC. Les premiers sont donnés par les modèles dynamiques : pour produire un certain type de note, un ingénieur se réfère généralement à des documents similaires, dont les préoccupations des auteurs sont supposées coïncider avec les siennes. Les critères de fond sont indiqués par les modèles de décomposition (objets, concepts), qui lui permettent d'accéder comme dans une liste à la configuration qui l'intéresse. En recoupant ces données, le système peut lancer une recherche d'information correspondant au besoin de l'utilisateur¹²⁶.

Production d'information : parallèlement à cette recherche d'information, un ingénieur a besoin de produire de l'information. La production d'information fait appel à d'autres applications supposées déjà à disposition de l'utilisateur (codes de calcul, traitement de textes, autres applications d'accès à l'information). Pour éviter les problèmes de transmission d'information entre applications, il est préférable d'utiliser des copier-coller et des glisser-déplacer, qui sont gérés par Windows.

4.3.1.2.2. Expert Chercheur

4.3.1.2.2.1. besoins

Les besoins en information des chercheurs diffèrent de ceux de l'ingénieur. Un chercheur peut avoir besoin d'explorer aussi profondément que possible toute l'information disponible sur un sujet particulier, puis passer à l'exploration poussée d'un autre sujet. Mais il peut également

avoir besoin de « passer » plusieurs calculs à l'aide de différents codes ou de différentes options, afin d'en comparer les résultats et finalement rédiger une note sans format prédéfini. Finalement, les chercheurs sont par définition les utilisateurs dont il est le plus difficile de prévoir les besoins par rapport au système d'information du SEC. Quoi qu'il en soit, nous pouvons décrire de manière générique leurs besoins en information par « trouver l'information la plus précise et la plus complète sur un domaine restreint et clairement identifié ».

4.3.1.2.2.2. caractéristiques particulières :

- bonne connaissance des paramètres et des mécanismes utilisés dans les calculs de criticité
- bonne connaissance des phénomènes physiques (neutroniques) liés à la criticité,
- utilisation détournée des outils (dans des buts pour lesquels ils n'ont pas été conçus),
- connaissance des domaines connexes.

4.3.1.2.2.3. Tâches

Navigation : la navigation entre les modèles est moins prévisible pour les chercheurs que pour les ingénieurs : elle ne suit pas un déroulement chronologique donné, mais s'articule autour d'un sujet précis (une configuration, un élément de code de calcul ou un type d'appareillage). Ce type de navigation « en rond » autour d'un point permet au chercheur de se faire une idée très précise du domaine limité qu'il explore.

Pour cela, le livre de connaissances doit faciliter ce type d'exploration en privilégiant les liens sémantiques inter-modèles et inter-sous-domaines plutôt que l'exploration séquentielle des diagrammes d'un même modèle ou d'un même sous-domaine.

Par exemple, pour un type de matière fissile donné, l'utilisateur peut vouloir explorer les diagrammes correspondant à la classification des matières fissiles (diagramme des concepts : classification ICSBEP), aux interactions physiques de cette matière avec d'autres corps (phénomènes), à l'utilisation qui en est faite dans les études de criticité (activités ou tâches) et à la façon dont elle est traitée dans les codes de calcul (activités). Pour les chercheurs, ce type de navigation se fait plus « par association » que pour les ingénieurs¹²⁷.

¹²⁶ Dans cette version du prototype, le modèle arborescent de décomposition du système d'information est censé remplacer le formulaire d'interrogation par champs. Nous avons depuis adopté des solutions plus « classiques ».

¹²⁷ Les types de navigation séquentiel et par association sont abordés au chapitre 5.

Recherche d'information : le chercheur est moins concerné que l'ingénieur par le format des éléments d'information qu'il recherche. L'essentiel pour lui réside dans le contenu de cet élément, car en général, l'utilisation qu'il en fait n'est pas de la recopie directe. Il peut cependant avoir besoin de spécifier un type d'élément d'information donné s'il sait où trouver ce qu'il cherche, mais les critères de forme restent pour lui un moyen de réaliser sa tâche de recherche d'information et non une partie de son besoin en information.

Du fait du modèle retenu pour la classification des informations, les fonctionnalités à mettre au point pour permettre la tâche de recherche d'information sont sensiblement les mêmes pour les chercheurs que pour les ingénieurs. Cependant, pour donner plus de poids aux critères de fond, les chercheurs se voient proposer plus d'espace pour expliciter et préciser leurs requêtes. De plus, leurs recherches se voulant plus exhaustives que celles des ingénieurs, les requêtes sont volontairement plus larges et génèreront plus de « bruit ».

Production d'information : les mêmes services sont ici à offrir aux chercheurs qu'aux ingénieurs. Les mécanismes de copier-coller et glisser déplacer leur offrent la possibilité d'utiliser le contenu des informations accédées pour la production de notes ou d'autres types d'informations.

4.3.1.2.3. Apprenant et novice

4.3.1.2.3.1. Besoins

Nous n'avons pas modélisé l'activité des utilisateurs apprenants ou novices en criticité ni n'en avons interviewé. Ce paragraphe se fonde donc sur les études de la littérature et sur l'argumentation présentées au paragraphe 4.2.4.1.

Les besoins de l'utilisateur apprenant sont dans un premier temps la découverte du domaine par l'exploration des diagrammes de connaissances et la navigation à travers des liens hypermédias explicitement typés. Ensuite, ses besoins évoluent progressivement vers ceux du type d'utilisateur auquel sa formation le destine.

4.3.1.2.3.2. Caractéristiques particulières :

- connaissance (supposée) des phénomènes physiques de base liés à la criticité,
- faible connaissance des méthodes de neutronique utilisés par les criticiens,
- peu ou pas de connaissances sur les outils de calcul de criticité
- peu ou pas de connaissances des activités du SEC,

- connaissance « normale » des outils de bureautique classiques (Word, Netscape, Acrobat Reader...).

4.3.1.2.3.3. Tâches

Navigation : quelqu'un qui utilise le livre de connaissances électronique pour découvrir la criticité débute avec le statut de novice. Il lui est donc proposé une vision d'ensemble du domaine grâce à la navigation à travers les diagrammes. La navigation dans le livre de connaissances hypermédia est ici le but principal de l'utilisateur. Les liens sont porteurs de sens ([Chaillot & Ermine 97]) et la navigation est d'autant plus enrichissante que l'utilisateur aura choisi et compris la relation qui lui permet de passer d'un diagramme à un autre.

Recherche d'information : pour les utilisateurs novices, des liens hypermédias statiques sont posés directement à partir des entités des diagrammes vers les documents les plus pertinents. Ces documents sont intégrés aux cartes de navigation, si bien que leur accès relève plus de la navigation à travers le livre de connaissances hypermédia que de la recherche d'information. Ce n'est que dans la phase d'apprentissage des outils en criticité que l'interface de recherche d'information devient active et possède les mêmes fonctionnalités que celle du type d'utilisateur que l'apprenant vise à devenir.

Production d'information : comme pour la recherche d'information, l'utilisateur novice n'a pas accès aux applications telles que les traitements de texte ou les codes de calcul. Ceci est justifié par le fait qu'il ne lui est pas demandé de production documentaire durant sa formation et qu'une utilisation correcte des codes nécessite un minimum de connaissances préalables en criticité.

4.3.1.2.4. Profils d'utilisation

Les tableaux ci-dessous récapitulent les caractéristiques de différents types d'utilisateurs ainsi que les fonctionnalités qui leur sont proposées par le prototype.

Utilisateur	Caractéristiques
Expert ingénieur	Bonne connaissance des paramètres et des mécanismes utilisés dans les calculs de criticité, ainsi que des outils à sa disposition, Connaissance moyenne des phénomènes physiques liés à la criticité, Connaissance de la terminologie technique et des références les plus utilisées,

	Pragmatisme (i.e. « time pressure »).
Expert chercheur	Bonne connaissance des paramètres et des mécanismes utilisés dans les calculs de criticité, - Bonne connaissance des phénomènes physiques liés à la criticité, Utilisation détournée des outils, Connaissance des domaines connexes.
Apprenant	Connaissances des phénomènes physiques de base liés à la criticité, Faible connaissance des méthodes de neutronique utilisées par les criticiens, Peu ou pas de connaissances sur les outils de calcul de criticité, Peu ou pas de connaissances des activités du SEC.
Novice	Aucune connaissance des phénomènes physiques de base liés à la criticité, Aucune connaissance des méthodes de neutronique utilisées par les criticiens, Aucune connaissance des activités du SEC.

Tableau 6. Caractéristiques des différents types d'utilisateurs en sûreté-criticité.

Sous-domaines	Novice	Apprenant	Expert
Études de criticité (ingénierie)	Navigation dans les diagrammes et les documents du SI les plus proches par des liens statiques	<ul style="list-style-type: none"> - Accès limité aux informations de type « ingénierie » - Accès aux options standard des codes 	<ul style="list-style-type: none"> - Accès « orienté-études » à l'ensemble du SI - Accès à toutes les applications
Expertises pour les transports (ingénierie)			<ul style="list-style-type: none"> - Accès « orienté-expertises » à l'ensemble du SI - Accès à toutes les applications
Codes de calcul (recherche)		<ul style="list-style-type: none"> - Accès limité aux informations de type « recherche » - Accès aux options standard des codes 	<ul style="list-style-type: none"> - Accès « orienté-codes » à l'ensemble du SI - Accès à toutes les applications
Expériences critiques (recherche)			<ul style="list-style-type: none"> - Accès « orienté-expériences » à l'ensemble du SI - Accès à toutes les applications

Tableau 7. Fonctionnalités offertes aux différents types d'utilisateurs.

4.3.1.3. Le modèle de l'information

Nous avons vu au chapitre 3 que le SEC possède un important fonds documentaire. La méthode de définition du modèle du système d'information est décrite au paragraphe 4.2.1.3. Le modèle de la partie « données » du système d'information réalisé est d'ailleurs présenté dans cette partie. Le diagramme suivant présente les outils qui sont disponibles au SEC.

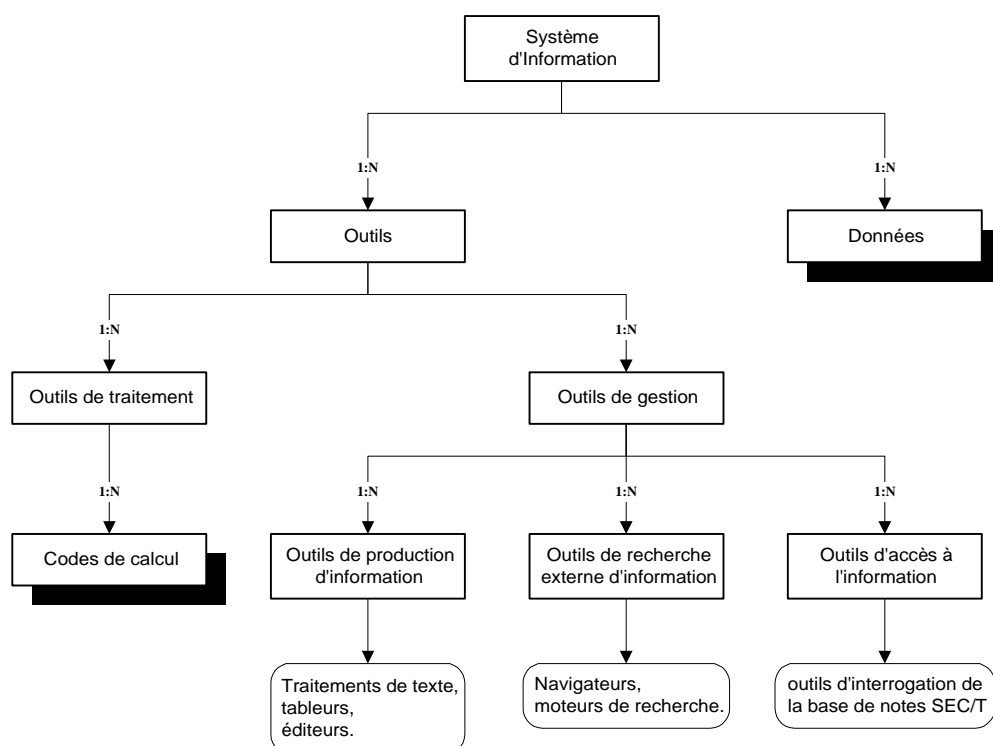


Figure 52. Les outils du système d'information du SEC.

4.3.2. Les outils conceptuels

La stratégie cumulative d'accès à l'information et le moteur de règles sont identiques à ceux décrits dans les paragraphes précédents. Nous ne les détaillons pas ici.

4.3.3. Fonctionnalités du système

Comme nous l'avons mentionné, les fonctionnalités initialement prévues sont, en plus de la navigation hypermédia et de l'accès à l'information, l'accès à tous les outils de production et de traitement de l'information disponibles au SEC. Nous avons abandonné cette idée pour deux raisons.

La première est que cette fonctionnalité est déjà assurée par le système d'exploitation. Notre objectif ici n'est pas de réaliser un « Windows bis » à l'intérieur d'un navigateur HTML. La communication entre les applications assurée par les liens OLE est suffisante pour le LCE. L'ouverture de ces applications peut être aisément réalisée par un lien hypermédia classique. Accéder à ces applications depuis la fenêtre du navigateur ne permettrait alors que de perdre le bénéfice du multi-fenêtrage.

La seconde raison concerne les codes de calcul, spécifiquement développés en collaboration avec des membres du SEC. L'inclusion de ces codes dans le LMCE aurait notamment conduit à la redéfinition de leur interface CRISTAL, développée au SEC. Or l'objectif du LMCE n'est pas d'interférer avec l'activité de l'organisation modélisée. Ce point, loin d'être anecdotique, pose la question de la validité des modèles après l'installation du LMCE dans l'organisation. Il est développé au chapitre 5.

4.3.4. État d'avancement du prototype

À la date de la fin de la thèse au CEA, les fonctionnalités de navigation dans les diagrammes de connaissance et d'accès à l'information à partir d'un élément d'un diagramme ont été implémentées dans le prototype de Jean-Marc Charlot. Ce démonstrateur a été présenté aux conférences ICNC'99 (International Conference on Nuclear Criticality-safety, cf. [Médini & al. 99]) et IAEA-SR-212/46 (conférence de l'Agence Internationale pour l'Énergie Atomique, cf. [Charlot & al. 99]). Il a d'ailleurs été très favorablement accueilli par les utilisateurs potentiels lors de ces présentations.

Ce démonstrateur utilise un parser java pour pré-transformer en HTML les fichiers XML contenant les diagrammes et les autres éléments visualisables. Ces éléments sont assemblés dans des cadres séparés et le fonctionnement dynamique de l'ensemble est assuré par des

fonctions JSCRIPT. De ce fait, le démonstrateur ne fonctionne qu'avec un navigateur Microsoft Internet Explorer 4.0 ou supérieur (seul ce type de navigateur est installé au SEC). Un exemple de l'interface de ce prototype est présenté ci-dessous.

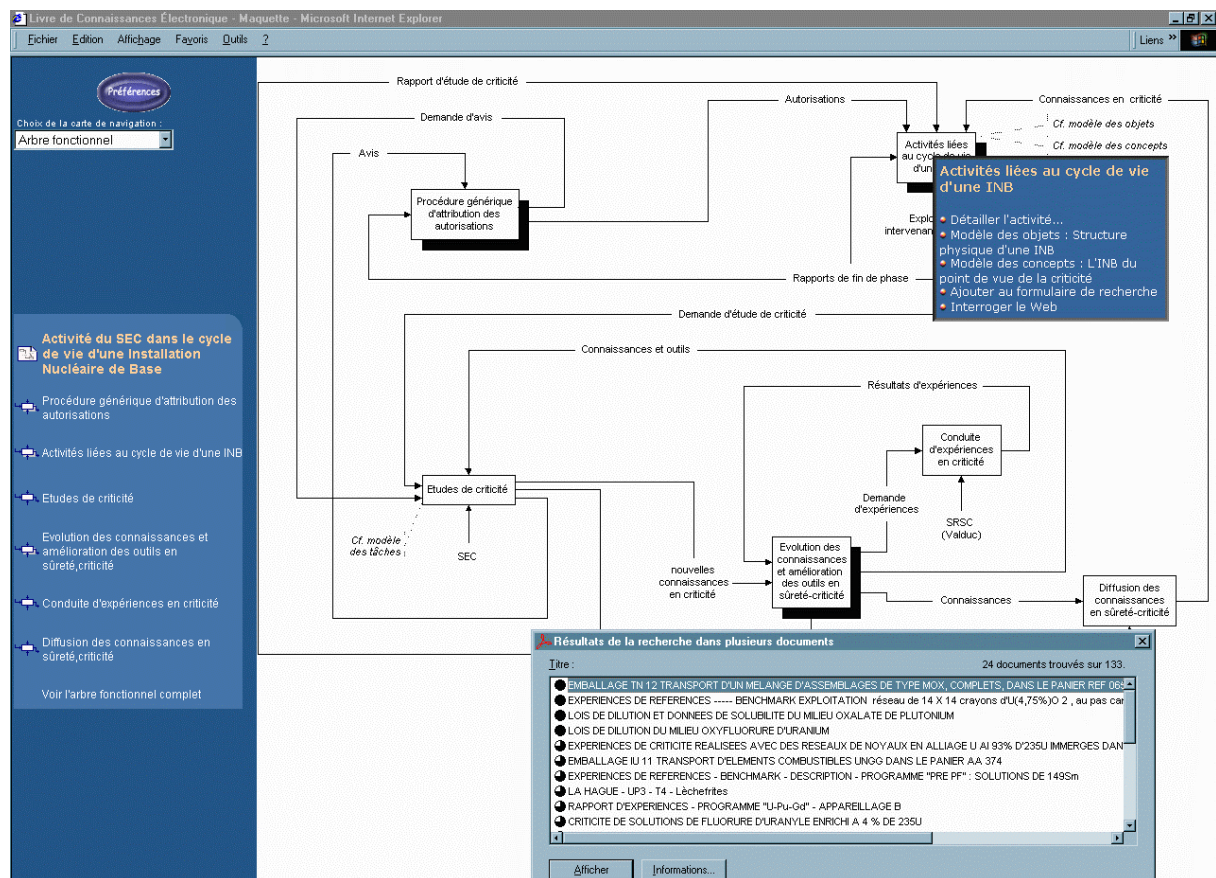


Figure 53 : exemple d'interface de navigation du prototype de LCE en sûreté-criticité nucléaire. La « combo box » située en haut à gauche permet de choisir d'accéder à un sous-domaine de la sûreté-criticité nucléaire, de visualiser un type de carte conceptuelle particulier ou d'accéder à la fonctionnalité de recherche d'information.

Jean-Marc Charlot a réalisé l'interface, la navigation et l'interrogation du moteur AltaVista® de ce prototype. Notre contribution a été de fournir les fichiers XML bien formés nécessaires à la visualisation des diagrammes de connaissances en sûreté-criticité et de permettre l'interrogation du fonds pdf via le contrôle ActiveX déclenchant notre plug-in Acrobat®. Nous avons mis cet outil en réseau au SIIA de la DIST et permis l'installation de ce contrôle ActiveX¹²⁸ sur des machines distantes appartenant au même réseau local que celle où il était disponible. La consultation du fonds pouvait alors être réalisée sur ces machines. Le but de cette opération était de tester la faisabilité technique de cette consultation. Nous n'avons pas poussé plus loin ces expérimentations pour deux raisons. Le public qui pouvait accéder à cette

¹²⁸ Via un fichier « cabinet » (.cab) lié à la page HTML contenant le formulaire.

application n'était pas le réel destinataire de l'outil et l'interface présentée n'était pas celle que nous voulions tester.

D'autre part, nous avons programmé un prototype de caddie à concepts (présenté en Figure 31) capable de sélectionner plusieurs éléments dans des diagrammes, de permettre la sélection des index et l'articulation de ces éléments et de construire la requête en langage VDK correspondante. Toutefois, l'interrogation à distance du fonds documentaire par le contrôle ActiveX d'appel du plug-in Acrobat® n'a plus été possible à partir de ce formulaire, du fait d'un changement intempestif des noms des machines réseau du SIIA. Nous n'avons pas modifié le code du contrôle ActiveX car la faisabilité de cette opération avait été montrée pour le formulaire d'interrogation précédent. De plus, les feuilles de style correspondant à ce caddie à concepts n'ont pas été développées, ce qui le rendait peu attrayant. Toutefois, ce prototype LCE permettant l'accès à l'information cumulatif fonctionnait en local.

Nous avons attendu la livraison d'un atelier de génération de LMCE¹²⁹ jusqu'à une semaine avant notre départ du CEA. C'est pourquoi la version définitive du LMCE tel qu'il est présenté ici n'est pas réalisée. Nous travaillons actuellement sans cet atelier et en utilisant les fonctionnalités du processeur Msxml. Il s'agit de mettre au point les fonctions JSCRIPT non implémentées et les feuilles de style XSL nécessaires à la visualisation du LMCE avec ce processeur.

4.4. Bilan de la conception

Dans ce chapitre, nous avons indiqué les différents aspects de la conception de notre prototype d'outil d'accès à l'information fondé sur la navigation hypermédia dans la modélisation de l'activité des utilisateurs. Cet outil repose sur trois modèles conceptuels qui permettent de donner un cadre formel à notre approche pluridisciplinaire d'accès à l'information. Une stratégie d'accès à l'information spécifique est également développée, dans la perspective de la prise en charge par le système de la formalisation des requêtes documentaires en fonction du contexte de travail des utilisateurs. La démarche de conception est ergonomique et s'appuie sur la modélisation effectuée dans l'unité lors de la phase de capitalisation des connaissances. La méthodologie de modélisation est adaptée au caractère pluridisciplinaire de l'approche et permet d'éviter les redondances dans la modélisation.

¹²⁹ Un projet est en cours avec la société C-Log™, qui développe actuellement un atelier de génération de LCE (MNÉSIK®), à partir du méta-atelier de Génie Lociciel GraphTalk®.

D'un point de vue logiciel, la démarche de conception est guidée par la définition du comportement de l'interface. Des critères de conception ergonomiques sont utilisés pour la conception et la mise en relation des différents composants de l'interface. Techniquement, notre prototype est conçu de façon à posséder une architecture réutilisable à travers les projets de gestion des connaissances menés à la DIST du CEA. Les éléments génériques et spécifiques à l'activité modélisée sont identifiés. Une architecture et des solutions techniques privilégiant l'utilisabilité parfois au détriment de la portabilité sont proposées.

Ce prototype, une fois évalué, servira de base à la discussion présentée au chapitre suivant. Celle-ci s'intéresse à la pertinence de la stratégie d'accès à l'information cumulative, et est ensuite élargie à l'évaluation de la pertinence de l'approche de la gestion des connaissances pour l'accès à l'information. Des adaptations de la méthode sont proposées et un bilan prospectif est dressé. Les conclusions tirées permettent d'adresser des perspectives d'évolution de l'outil et de l'approche d'accès à l'information par recontextualisation.

Bibliographie

- [Adobe 99] : Adobe Developer Technologies, "*Highlight File Format*", Technical Note #5172, Adobe Systems Inc., date de dernière révision de la version utilisée : 10 nov. 1999, disponible en ligne : <http://partners.adobe.com/asn/developer/technotes/TN5172.pdf>.
- [Agosti & Smeaton 96] : "*Information Retrieval and Hypertext*", Agosti M., Smeaton A. Eds., Kluwer Academic Publications, 1996.
- [Asse 85] : Asse A., "*Aide au diagnostic industriel par des méthodes basées sur la théorie des sous-ensembles flous*", thèse de doctorat de troisième cycle, Université de Valenciennes, avril 1985.
- [Bastien & al. 98] : Bastien, J. M. C., Leulier C., Scapin D-L, "*L'Ergonomie des sites web*", in *Créer et Maintenir un service web*, J-C. Le Moal & B. Hidoine Eds., ADBS, Paris, pp. 111-173, 1998.
- [Chaillot & Ermine 97] : Chaillot M., Ermine J-L., "*Le livre de connaissances électronique*", Document numérique, 1(1), 75-98, 1997.
- [Charlot & al. 00] : Charlot J-M., Chaillot M., Médini L., "*Information Engineering based on knowledge modelling: a case study in nuclear criticality safety*", International Seminar on Integrated Information Systems, IAEA-SR-212/46, Vienne (Austria), avril, 10-14, 2000.
- [Fluhr & al. 97] : Fluhr C., Schmit D., Ortet P., Elkateb F., Gurtner K., "*SPIRIT-W3: a distributed crosslingual indexing and search engine*", in proceedings of the 7th Annual Conference of the Internet Society INET'97, the Internet, the Global Frontier, Juin 97.
- [Hunter & al. 01] : Hunter D., Cagle C., Gibbons D., Ozu N., Pinnock J., Spencer P., "*Initiation à XML. Avec trois études de cas détaillées*", Traduit par F. Lemainque, L. Adam, C. Raspauld, Eyrolles, Paris 2001.
- [Kolski 97] : Kolski C., "*Interfaces homme-machine, application aux systèmes industriels complexes*", 2^{ème} édition revue et augmentée, Hermès, Paris, 1997.
- [Lainé-Cruzel 94] : Lainé-Cruzel S., "*Vers de nouveaux systèmes d'information prenant en compte le profil des utilisateurs*", Documentaliste, Sciences de l'information, vol. 31, No. 3, pp. 143-147, 1996.
- [Lainé-Cruzel & al. 96] : Lainé-Cruzel S., Lafouge T., Lardy J-P. & Ben Abdallah N., "*Improving information retrieval by combining user profile and document segmentation*", Information Processing and Management, Elsevier Science Ltd, Vol. 32, No. 3, pp. 305-315, 1996.
- [Mackay & al. 98] : Mackay W-E., Fayard A-L., Frobert L., Médini L., "*Reinventing the familiar: exploring an augmented reality design space for air traffic control*", Proceedings of CHI'98, Conference on Human Factors in Computing Systems, Los Angeles, pp. 558-565, Avril 1998.
- [Marchionini 95] : Marchionini G., "*Information seeking in electronic environments*", New York, Cambridge University Press, 1995.
- [Médini 97] : Médini L., "*Conception et Réalisation d'un Livre de Connaissances Électronique sous forme de Document Hypermédia*", Rapport de projet de DEA, Orsay et INSTN, Sept. 1997.
- [Médini & al. 99] : Médini L., J. Anno, M. Chaillot, V. Rouyer, D. L. Scapin, "*Towards a knowledge oriented system for information management in nuclear safety*", Proceedings of

- the Sixth International Conference on Nuclear Criticality-Safety, Versailles, Sept. 1999, 1451-1461.
- [Médini 99] : Médini L., *"Utilisation de la modélisation des connaissances pour la conception d'un référentiel hypermédia d'accès à l'information scientifique et technique"*, Rencontres Doctorales, Actes de la conférence IHM'99, vol. 2, pp. 174-177, Montpellier, Nov. 1999.
- [Nielsen 90] : Nielsen J., *"Hypertext & Hypermedia"*, Academic Press Inc., San Diego, Ca., 1990.
- [Palmquist & Kim 98] : Palmquist R. A., Kim K-S., *"Modeling the Users of Information Systems: Some Theories and Methods"*, Electronic Resources: Use and User Behavior, Hemalata Iyer (Ed.), The Haworth Press Inc., pp. 3-25, 1998.
- [Pardi 99] Pardi W-J., *"XML en Action"*, Microsoft Press, 1999.
- [Roger & al. 98] : Roger D., Lavandier J., Kolmayer E., *"Navigation et interfaces : cartes conceptuelles et autres outils"*, analyse bibliographique réalisée pour PARINFO (Programme d'Aide à la Recherche en INFormation), révision du 27 Avril 2000, disponible en ligne : <http://www.enssib.fr/bibliotheque/documents/travaux/kolmayer/kolm-navig0.html>.
- [Rouet 98] : Rouet J-F., Introduction aux PréActes du Quatrième Colloque Hypermédias et Apprentissages, Poitiers, 15-17 octobre 1998.
- [Rouet & Tricot 95], Rouet J-F., Tricot A., *"Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive"*, Sciences et techniques éducatives, vol. 2, n° 3, Hermès, Paris, 1995.
- [Rouet & Tricot 98] : Rouet J-F., Tricot A., *"Chercher de l'information dans un hypertexte: vers un modèle des processus cognitifs"*, in Les hypermédias, approches cognitives et ergonomiques, Hermès, Paris, pp. 57-74, 1998.
- [Tricot & al. 98] : Tricot A., Pierre-Demarcy C., El Boussarghini R., *"Définition d'aides en fonction des types d'apprentissages dans les environnements hypermédia"*, in Quatrième Colloque Hypermedias et Apprentissages, Poitiers, pp. 41-58, 15-17 octobre 1998.
- [Waern & Hägglund 97] : Wærn, Y., Hägglund S., *"User Aspects Of Knowledge-Based Systems"*, Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu (eds.), Elsevier Science, 1997, pp. 1159-1176.
- [Woodhead 90] : Woodhead N., *"Hypertext & Hypermedia. Theory and Applications"*, Sigma Press, GB, 1991.

Chapitre 5. Discussion : l'apport et les perspectives de notre approche pour la problématique de l'accès à l'information

Le chapitre précédent détaille l'approche et les aspects techniques de la contribution de cette thèse. Il s'agit ici d'en décrire les apports méthodologiques et de tenter d'en évaluer l'intérêt pour la problématique de l'accès à l'information. Pour cela, ce chapitre s'organise en deux parties. Tout d'abord, nous soulignons ce qui constitue à notre sens les apports et les limites méthodologiques de ce travail. Puis nous proposons un bilan prospectif dans lequel nous distinguons la suite immédiate de ce travail (i.e. la validation des résultats obtenus) des perspectives de prolongement de celui-ci.

5.1. Spécificités de l'approche

Dans cette partie, il s'agit de mettre en lumière les apports et les limites du travail présenté aux chapitres précédents. Dans une première sous-partie, nous présentons les spécificités de notre approche par rapport à la problématique de l'accès à l'information. Il s'agit d'aborder les aspects méthodologiques de notre démarche pluridisciplinaire. La deuxième sous-partie s'intéresse aux aspects ergonomiques de l'outil réalisé. Les apports de notre approche sont discutés sous l'angle de leur impact sur l'interface et le comportement de l'utilisateur. La troisième sous-partie aborde l'ingénierie de l'information. Plus précisément, l'intérêt de l'indexation conceptuelle est discuté dans le cadre de notre démarche d'accès à l'information par recontextualisation. La dernière sous-partie de ce paragraphe est consacrée aux apports méthodologiques spécifiques à la méthode MKSM et au domaine de la gestion des connaissances. Cette discipline est considérée dans les cadres de l'accès à l'information, de la conception logicielle et de sa ré-injection dans l'activité qu'elle modélise.

5.1.1. Aspects méthodologiques

La principale spécificité de la démarche d'accès à l'information présentée dans ce travail est sa pluridisciplinarité. Elle fait appel à la gestion des connaissances, à l'ergonomie et aux sciences de l'information (i.e. conception hypermédia, indexation et recherche d'information) pour mettre au point les trois modèles conceptuels sur lesquels s'appuie notre approche. Sans prétendre à l'exhaustivité, cette approche tridimensionnelle nous semble relativement

complète et permet de dégager un point de vue systémique de l'accès à l'information que nous n'avons pas trouvé dans la littérature jusqu'à présent.

En revanche, les travers inhérents à la complétude de l'approche sont bien entendu la lourdeur de la phase de modélisation nécessaire à sa mise en place et la complexité du système réalisé qui implique un effort de formation à l'outil de la part de l'utilisateur. Nous pensons d'ailleurs accompagner le prototype d'un manuel d'utilisation et d'une « hotline », sous la forme d'une adresse e-mail à laquelle les utilisateurs pourront envoyer les questions qu'ils se poseront sur l'utilisation du prototype.

Par ailleurs, si l'on peut considérer comme un apport d'avoir positionné et mis en relation ces disciplines dans le cadre de l'étude d'une démarche spécifique d'accès à l'information, la principale limite de notre approche est la difficulté d'approfondir suffisamment tous les domaines abordés. En particulier, le niveau de détail des modélisations obtenues peut paraître plus superficiel que dans certaines applications spécifiques à chacune des disciplines abordées et les techniques employées moins innovantes.

D'un point de vue plus spécifique, cette approche s'appuie sur deux paradigmes : l'accès à l'information par recontextualisation et la stratégie cumulative. Dans les paragraphes suivants, nous détaillons successivement leurs particularités.

5.1.1.1. Avantages et limites de la démarche de recontextualisation de l'information.

Le fait de modéliser et de proposer une navigation hypermédia dans le contexte d'utilisation d'un système d'information plutôt que dans ce système lui-même n'est pas nouveau. L'exemple le plus souvent cité de ce paradigme est la métaphore spatiale utilisée dans les OPAC (Online Public Access Catalogues) américains pour aider les utilisateurs à trouver plus facilement l'information qu'ils recherchent. Dans ces systèmes, l'utilisateur se déplace dans une bibliothèque virtuelle, qui peut être vue comme le contexte commun de l'accès à l'information pour tous les lecteurs. Dans d'autres domaines que celui de l'accès à l'information, nous pouvons citer l'exemple des « assistants » hypermédias, qui guident les utilisateurs le long des différentes étapes de réalisation d'une tâche informatisée.

Plus généralement, les logiciels « orientés-utilisateurs » tendent à réduire les gaps qui séparent les problèmes que se posent les individus des tâches calculatoires qui permettent de les solutionner. Pour concevoir un tel système, il est nécessaire d'appréhender à la fois la complexité du problème posé à l'utilisateur et des méthodes employées par celui-ci pour le

résoudre. Ces deux espaces – problème et résolution – sont par exemple modélisés dans les systèmes issus de l'Intelligence Artificielle, comme les systèmes experts destinés à supplanter la résolution de problèmes par l'opérateur humain. Cette approche a montré ses limites en termes d'opérationnalisation. Cependant, elle donne de bons résultats en matière de modélisation.

Dans le domaine de l'accès à l'information, le but de l'indexation conceptuelle est de s'approcher autant que possible de la vision de l'utilisateur et de sa problématique d'accès à l'information. Il s'agit ensuite de mettre en œuvre un modèle de l'espace résolution qui permet de faire correspondre cette description en termes de problème avec une indexation structurelle des sources d'information à partir de laquelle est menée la tâche de recherche d'information.

La modélisation de l'activité de l'utilisateur permet d'obtenir et de représenter ces deux espaces. Les taxinomies modélisent la façon dont l'utilisateur perçoit son problème et les actinomies celle dont il le résout (voir définitions, chapitre 3). Notre approche d'accès à l'information se fonde sur la navigation hypermédia à travers ces deux espaces pour assister l'utilisateur dans la traversée du gap qui le sépare de l'envoi de sa requête à un moteur de recherche d'information.

L'originalité de cette démarche est qu'elle est complète sans être impérative. Elle est complète car elle modélise la totalité de l'activité de l'utilisateur. En se fondant sur les points de vue d'experts du domaine, elle décrit une vision aussi complète que possible des espaces problème et résolution. Cependant, contrairement à un guidage linéaire de l'utilisateur dans sa stratégie de résolution de son problème d'information, elle ne propose qu'un référentiel de navigation à travers les différentes étapes de cette résolution. L'individu confronté à un problème d'information peut naviguer librement dans ces deux espaces. Nous espérons ainsi – mais cela reste à prouver – qu'elle est adaptable à un grand nombre de variations en termes de besoins d'information et de spécificités des utilisateurs.

5.1.1.2. Les stratégies d'accès à l'information proposées par le LMCE

D'un point de vue général, le LMCE possède la particularité de proposer trois stratégies d'accès à l'information distinctes en fonction des différents types d'utilisateurs répertoriés. Une vue schématique de cet outil en termes de stratégies d'accès à l'information, est présentée ci-dessous.

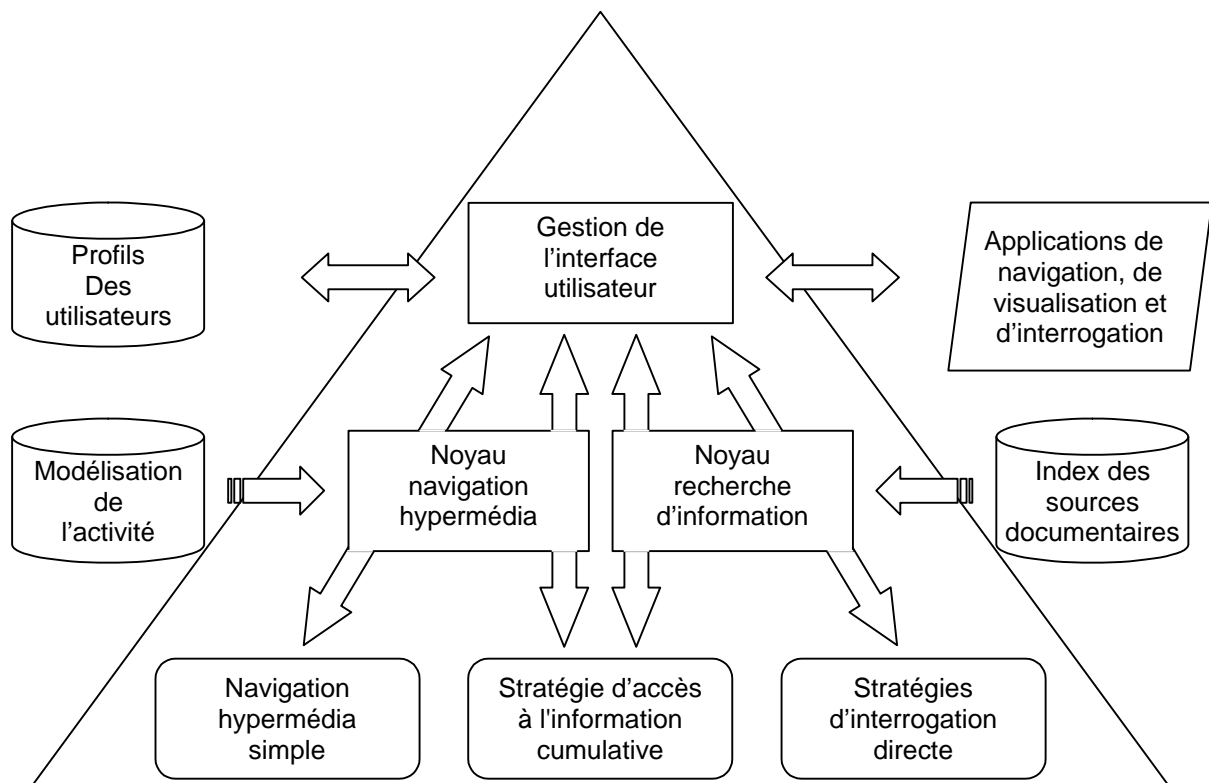


Figure 54. Architecture conceptuelle du LMCE.

Cette architecture comporte trois niveaux. Un premier niveau « interface » gère les fonctionnalités globales du système, telles que la connexion des utilisateurs et le choix de la stratégie d'accès à l'information à mettre en œuvre. Il interagit depuis le navigateur W3 avec la base des profils utilisateurs, les outils de recherche d'information et de visualisation des documents. Techniquement, cette interaction fait en grande partie appel à la technologie ActiveX pour l'écriture dans les fichiers ou la prise en charge des API des outils ci-dessus.

Le deuxième niveau est un niveau « générique ». Il prend en charge séparément les fonctionnalités des stratégies de parcours et analytiques. La prise en charge des stratégies de parcours est réalisée d'après un modèle hypermédia hiérarchique, qui décompose l'hyperdocument proposé en sous-domaines, formalismes, diagrammes et nœuds. Le typage des liens proposés est défini en conséquence : ceux-ci peuvent être intra-diagramme, inter-diagrammes, inter-modèles, inter-sous-domaines (voir [Médini 97]) ou externes (cf. les liens statiques du LCE de [Chaillot & Ermine 97]). L'envoi du contenu textuel des nœuds pour la recherche d'information est ici considéré comme un lien externe, qui est traité dans l'autre partie de ce niveau. Techniquement, l'association XML/XSL (et CSS) pour les diagrammes, HTML/CSS pour les cartes conceptuelles et JScript pour le traitement des liens suffisent à obtenir la présentation dynamique de l'activité de l'utilisateur recherchée.

En parallèle de la gestion de la partie hypermédia, le niveau générique du LMCE assure la prise en charge de l'interrogation. Il reçoit du niveau supérieur les caractéristiques de l'utilisateur (essentiellement, la donnée du filtrage à appliquer) et du niveau inférieur, le besoin d'information (i. e. : la requête). Cet élément se charge alors de transformer ces données comme cela est détaillé au paragraphe 4.2.3.2. Techniquement, il procède par manipulation de chaînes de caractères pour définir les requêtes à envoyer aux différents index qui constituent le modèle de l'information. Cette manipulation est actuellement réalisée par des fonctions JScript.

Le troisième niveau, que nous appellerons « spécifique », représente les trois stratégies d'accès à l'information proposées par le LMCE. Chacune d'elle prend en charge les différentes fonctionnalités qui lui ont été attribuées au chapitre 4 et instancie l'un ou les deux élément(s) du niveau générique. À ce niveau est éventuellement gérée la formalisation de la requête (en fonction du type de formulaire d'interrogation).

La navigation hypermédia et les stratégies analytiques ne sont pas discutées dans la suite de ce paragraphe. En effet, la deuxième spécificité de notre approche est la stratégie d'accès à l'information cumulative. Cette stratégie est conçue dans le même esprit que le parcours libre des espaces problème et résolution définis au paragraphe précédent : il s'agit de proposer à l'utilisateur un support où il peut stocker et organiser les caractéristiques de son besoin d'information et les traiter de façon non linéaire. Nous avons choisi de proposer cette solution car nous supposons que la séparation des tâches de sélection des éléments du besoin d'information et d'articulation de ces éléments peut réduire la charge cognitive de l'utilisateur lors de la composition de sa requête. Si tel n'est pas le cas, le formulaire proposé permet de réaliser ces deux tâches en parallèle, comme avec un formulaire « linéaire ».

5.1.2. *Aspects ergonomiques de notre démarche*

Du point de vue de l'interaction homme-machine, il convient de discuter plus en détail les modalités de la mise en œuvre de notre démarche d'accès à l'information par recontextualisation. En effet, l'hypothèse sur laquelle se fonde cette thèse est que les diagrammes de connaissances présentent les éléments supposés correspondre aux caractéristiques d'un besoin d'information dans le contexte de l'accomplissement d'une tâche modélisée. Le bien-fondé de notre stratégie d'accès à l'information repose sur l'accessibilité de toutes les composantes de ce besoin d'information.

Trois éléments contribuent à cette accessibilité. Il s'agit d'une part de s'assurer que les diagrammes comportent effectivement toutes les caractéristiques des besoins d'information qui peuvent se poser aux utilisateurs lors de l'accomplissement de l'activité modélisée. D'autre part, il s'agit de déterminer si les diagrammes permettent effectivement aux individus d'accéder à ces caractéristiques et de les identifier. Enfin, il s'agit d'évaluer l'intérêt de la sélection d'éléments des diagrammes (correspondant ou non à des critères d'indexation documentaire) par rapport à l'expression de ces caractéristiques grâce à un autre mode d'interaction (en langage naturel ou formel par exemple).

5.1.2.1. Présence des caractéristiques du besoin d'information dans les diagrammes

L'exhaustivité de la collection de diagrammes en termes de caractéristiques des besoins d'information des différents types d'utilisateurs repose sur la qualité et la profondeur de la modélisation des connaissances réalisée. Nous indiquons notamment au chapitre 4 les spécificités de cette modélisation dans le cadre de la conception d'un LMCE.

Nous pouvons remarquer que la qualité du filtrage documentaire dépend de l'adéquation de la tâche de l'utilisateur par rapport à la modélisation. En d'autres termes, il s'agit de définir le niveau de précision nécessaire et suffisant pour la prise en compte des caractéristiques de chacune des tâches de l'activité modélisée impliquant un accès à l'information. Une modélisation effectuée en dessous de ce niveau de précision risque de ne pas permettre un accès à l'information pertinent pour tous les problèmes d'information rencontrés. Un niveau de profondeur trop important dans la modélisation entraîne d'une part un surcoût du projet de gestion des connaissances et d'autre part un risque accru de désorientation de l'utilisateur dans l'hypermédia.

5.1.2.2. Accessibilité des caractéristiques du besoin d'information par la navigation hypermédia dans les diagrammes

Seule une évaluation ergonomique du prototype peut apporter une réponse objective à l'utilisabilité de cet outil. Les résultats de cette évaluation doivent permettre d'apprécier l'efficacité de la tâche de sélection des éléments des diagrammes correspondant aux différentes caractéristiques des besoins d'information des utilisateurs.

En l'absence de résultats sur ce point, nous pouvons toutefois nous référer aux travaux cités au chapitre 2 (paragraphes 2.2.7 et 2.2.8), qui présentent des outils proposant des stratégies d'interrogation documentaire à partir de fragments d'hypertextes. Parmi ces travaux, les plus

proches de notre problématique sont ceux de [Golovchinsky & Chignell 93] et de [Zizi & Beaudoin-Lafon 94].

Les premiers permettent la formulation graphique de requêtes booléennes complexes. Une évaluation de l'efficacité du système réalisé montre l'intérêt de cette stratégie en termes de pertinence de la recherche d'information. L'utilisabilité du système n'est cependant pas évaluée. Ces travaux ne sont pas suffisamment nombreux pour nous permettre de généraliser les conclusions tirées de ces travaux en une évaluation de l'intérêt de cette stratégie au LMCE. Leur existence et les conclusions tirées nous confortent toutefois dans l'idée que notre stratégie d'accès à l'information cumulative peut être efficace si son utilisabilité est avérée.

Les seconds permettent la formulation graphique de requêtes documentaires à partir d'éléments des cartes conceptuelles situées « au-dessus » du corpus d'information, au même titre que les diagrammes de connaissances. En effet, celle-ci est considérée comme un référentiel permettant d'accéder aux éléments du système d'information. La collection de diagrammes a alors le même statut qu'une carte conceptuelle de ce système d'information¹³⁰. Dans ces conditions, l'identification dans les diagrammes des caractéristiques d'un besoin d'information peut être assimilée à une tâche de recherche d'information dans l'hypertexte constitué de la collection de diagrammes.

Cette tâche de recherche d'information dans la collection de diagrammes s'ajoute à la tâche principale d'accès à l'information dans le cadre de l'activité de l'individu et à cette dernière activité. L'accès à l'information par recontextualisation est donc une charge cognitive supplémentaire pour l'utilisateur, au même titre qu'une quelconque autre tâche de formalisation du besoin d'information. Pour J-F. Rouet et A. Tricot ([Rouet & Tricot 98]), cette surcharge s'accumule dans la mémoire de travail de l'individu, qui possède une capacité limitée. Le problème est de savoir si cette troisième sous-tâche est plus ou moins contraignante pour l'individu que la formalisation d'un besoin d'information par un autre moyen.

Un des désagréments prévisibles des utilisateurs du LMCE risque d'être la difficulté de se construire une stratégie de parcours générique cohérente dans l'hypertexte. En effet, pour ne pas avoir l'impression de perdre du temps à chercher les éléments des diagrammes à inclure dans leurs requêtes, ils ont besoin de savoir « où » sont les éléments pertinents des

¹³⁰ Dans ces conditions, il est intéressant de noter que les cartes conceptuelles spécifiques à la méthode MKSM (arbre fonctionnel, dictionnaire des concepts...) sont un troisième niveau conceptuel d'aide pour l'accès à l'information. Elles permettent l'accès aux diagrammes qui permettent l'accès à l'information.

diagrammes. Pour cela, il convient de former ces individus à l'utilisation de l'outil. L'objectif est notamment de les aider à identifier la position des différents types de critères d'indexation dans les différents formalismes MKSM. Une certaine connaissance du contenu de la modélisation réalisée et des liens hypermédias disponibles est à notre avis recommandée pour arriver à atteindre relativement rapidement un élément d'un diagramme.

5.1.2.3. Sélection d'éléments correspondant à des critères d'indexation

Nous avons vu au chapitre 4 que le choix de présentation d'un formalisme correspondant à un type de critère d'indexation est déterminant pour l'utilisabilité de notre approche d'accès à l'information cumulative. Les critères d'indexation concernent a priori les éléments du domaine sur lesquels il est le plus pertinent de mener une recherche d'information. C'est pourquoi il est utile de pouvoir les identifier rapidement dans les diagrammes. Ce paragraphe identifie les formalismes concernés par la sélection d'éléments correspondant aux différents critères.

5.1.2.3.1. *Identification des critères de forme*

L'identification de ces critères se fait grâce aux modèles dynamiques. Le modèle des activités permet de disposer d'une vue globale de l'activité synchronique de la partie de l'organisation à laquelle l'utilisateur s'intéresse. Quel que soit ce niveau de granularité, cette activité est positionnée dans un contexte global en soulignant l'ensemble des flux grâce auxquels elle est réalisée.

Cet aspect global de la modélisation de l'activité permet d'identifier en priorité les critères formels du besoin d'information. Les types de flux documentaires sont notamment importants pour l'accès à l'information cumulatif. Les critères de forme identifiés comme spécifiques à certains types d'éléments des diagrammes du modèle d'activité sont les critères « auteur » ou « origine » et « type de document ». Ces critères correspondent respectivement aux éléments de type « acteur » et « flot ou ressource » du modèle des activités¹³¹.

Les modèles dynamiques diachroniques d'évolution de [Van Craynest & al. 00], s'ils ne sont pas utilisés dans le projet-pilote, sont tout de même significatifs pour l'identification des

¹³¹ Dans l'optique de l'accès à l'information, il serait utile de distinguer deux types de flux et de ressources que le formalisme des activités ne distingue pas. Il s'agit des flots matériels des flots documentaires issus des activités. Seuls ces derniers nous intéressent ici.

critères de forme. Ces formalismes comportent des informations temporelles permettant d'identifier les dates de rédaction des documents recherchés.

5.1.2.3.2. *Identification des critères de fonds*

Nous avons vu plus haut qu'une notion définie dans un diagramme de concepts est choisie comme référence pour l'indexation conceptuelle. Le diagramme correspondant est donc à la source de l'identification par l'utilisateur des critères de fond pertinents pour sa recherche.

5.1.2.4. *Sélection des autres éléments des diagrammes*

Des éléments ne correspondant pas à des critères d'indexation sont disponibles a priori dans tous les diagrammes, excepté le diagramme des concepts définissant les critères de fond. Les formalismes MKSM présentent différents types d'éléments dont la sélection en tant qu'élément du besoin d'information correspond à des *valeurs sémantiques* différentes, en fonction desquelles sont définis des *liens sémantiques* différents. Deux types de valeurs sémantiques sont identifiés.

5.1.2.4.1. *Éléments de contexte*

Ces éléments sont destinés à permettre à l'utilisateur de définir le contexte de sa recherche. Il ne spécifie pas là le corps de sa requête mais l'activité ou la tâche qui préside à sa démarche d'accès à l'information. Ce sont donc ces formalismes dynamiques synchroniques qui sont concernés par la sélection de ces éléments. La sélection d'une activité correspond alors à l'identification par l'utilisateur d'un sous-domaine d'activité, des types de flux documentaires qui y sont rattachés et éventuellement d'un auteur ou d'une organisation d'origine des documents recherchés. Le même type d'inférence peut être réalisé à partir d'une tâche, s'il est possible d'identifier l'activité dans laquelle elle s'inscrit.

Techniquement, un *lien contextuel* est posé sur l'élément en indiquant dans l'élément XML du diagramme correspondant une valeur de reformulation comportant l'ensemble de ces caractéristiques. Il est à noter que ce type de lien n'est pas exclusif. Un autre type de lien sémantique peut coexister dans la balise.

5.1.2.4.2. *Éléments de requête*

Tous les éléments des diagrammes peuvent donner lieu à une interrogation sur leur texte. Dans ce cas, la valeur sémantique de ces éléments n'est autre que de constituer le corps de la

requête. La valeur de reformulation de ces éléments est donc uniquement fonction de leur texte.

5.1.2.4.3. *Intérêt des autres formalismes*

D'un point de vue diagrammatique, la structure de ces modèles est déterminante pour leur pertinence en matière d'accès à l'information. Par exemple, les formalismes arborescents (i.e. les objets) ou présentés sous forme de listes (les tableaux de texte) peuvent être utilisés pour présenter des critères d'indexation de fonds, comme c'est le cas dans notre projet-pilote avec le modèle des concepts traitant des différents types de matières fissiles. Les modèles de flots de données peuvent présenter les critères de forme et les modèles d'historique peuvent identifier des dates ou des périodes correspondant aux dates de rédaction des documents recherchés. Il convient alors de définir, pour les éléments pertinents de ces modèles, les valeurs et les liens sémantiques correspondants.

5.1.2.5. Bilan ergonomique de l'intérêt de la modélisation des connaissances comme moyen de présentation hypermédia des caractéristiques du besoin d'information

Les conclusions des paragraphes précédents nous confortent dans l'idée de la pertinence d'une stratégie d'accès à l'information fondée sur la navigation hypermédia et la sélection des composantes du besoin d'information. Pour les individus non habitués l'utilisation des outils électroniques de recherche d'information, cette démarche permet l'articulation des composantes du besoin d'information et prend en charge la formalisation de la requête. Toutefois, son principal inconvénient est de nécessiter une sous-tâche de recherche d'information dans l'hypermédia pour identifier les éléments pertinents du besoin d'information. Dans le cas d'un livre de connaissances hypermédia, cette sous-tâche peut être complexe, du fait de son organisation a priori non conçue pour l'accès à l'information.

Ici, la problématique de la conception d'un hypermédia d'accès à l'information s'oppose donc à celle de la conception du livre de connaissances. [Tricot & Nanard 98] (p. 48) se réfèrent à [De Vriès & de Jong 97] et [Bernstein 93] pour affirmer que le fait de vouloir accomplir une activité avec un système prévu pour une autre est voué à l'échec¹³². L'adéquation de l'outil par rapport à la finalité d'accès à l'information est donc essentielle pour l'efficacité de l'accès à l'information. L'objectif de ce paragraphe est de comparer la pertinence de la structuration

de ce LCH pour l'accès à l'information avec celle d'hypermédias spécifiquement conçus pour cette finalité.

De tels hypermédias sont par exemple les « rubriques » proposées par certains moteurs de recherche sur le web (cf. par exemple le moteur de recherche Yahoo®, <http://www.yahoo.fr>) ou la généralisation des « sommaires » hypermédias dans les aides à l'utilisation des logiciels (cf. les dernières versions de la suite Microsoft Office®). Les cartes conceptuelles (ou « site maps ») présentes sur de nombreux sites web sont un autre exemple de ces hypermédias dédiés à la recherche d'information. Le point commun de ces exemples est qu'ils sont généralement structurés de façon arborescente. Les travaux de [Crampes & Ranwez 00] donnent un cadre plus formel à cette généralisation. Ces auteurs s'appuient sur la notion d'ontologie pour définir des styles de navigation conceptuels sur le web. Plus généralement, les travaux de thèse de Sylvie Ranwez s'appuient sur les ontologies pour la structuration d'hypertextes [Ranwez 00]. Cependant, cet auteur limite sa définition des ontologies à des structures arborescentes et ne tient pas compte des autres types de liens sémantiques pouvant exister entre les différents éléments de ce type de classification. En ce sens, il s'agit plutôt de taxinomies au sens de [Vogel 88], c'est-à-dire en tant que types de classifications statiques.

Quoi qu'il en soit, pour cet auteur, « Le savoir naturel est d'abord classificatoire. (...et...) L'activité de classification représente la condition minimale de l'adaptation » (p. 99). Cela explique à notre sens l'émergence de ce type de structures arborescentes dans les tables des matières et autres types de cartes conceptuelles de structuration des corpus documentaires. C'est également la raison pour laquelle notre modèle du système d'information présenté au chapitre 4 est arborescent.

Cependant, si pour Claude Vogel, les taxinomies suffisent pour aboutir à la structure des bases de connaissances vues par les experts, la méthode MKSM s'appuie sur des formalismes plus complexes (i.e. des actinomies) exploitant l'ensemble de l'activité des experts. La mise au format hypermédia de cette représentation implique donc des types de liens aux sémantiques différentes (cf. [Médini 97]). L'utilisateur du LMCE cherchant à atteindre un item depuis un diagramme où il ne figure pas doit donc envisager toutes les inférences possibles à partir des éléments du diagramme courant, puis les inférences possibles à partir de tous les éléments de chacun des diagrammes cibles etc.

¹³² La gestion des connaissances a d'ailleurs prouvé dans ce domaine qu'elle n'est pas aisément applicable à autre chose qu'à la transmission des savoir-faire dans l'organisation. Voir chapitre 3, à propos de l'utilisation de la méthode MKSM pour l'opérationnalisation, l'aide à la décision ou les codes de calculs scientifiques.

À notre sens, la différence de complexité entre la sélection d'items dans une taxinomie et dans une collection de diagrammes MKSM est comparable à celle existant entre les jeux de dames et d'échecs. Nous assimilons ici les transitions (liens hypermédias) entre les différents états de l'interface (page écran de l'hypermédia présentée) aux transitions (mouvements des pièces) entre les états (positions des pièces) du jeu. Dans le premier jeu, comme dans le cas des taxinomies, les états du système sont constitués d'une grande partie d'éléments similaires et chaque type d'éléments de ces systèmes permet un nombre limité (un ou deux) de types de transitions¹³³. Un échiquier complet ou un écran présentant un diagramme MKSM comporte de nombreux types d'éléments différents, à partir desquels plusieurs types de transitions sont possibles. L'apprentissage et la pratique sont donc moins complexes pour le jeu de dames que pour le jeu d'échecs. Il en va de même de la prise en main et de l'utilisation comparatives d'un système d'accès à l'information proposant une carte conceptuelle du corpus et du LMCE. Du point de vue de la complexité des stratégies mises en œuvre, le LMCE apparaît donc comme un outil plus lourd à manipuler pour l'utilisateur que les taxinomies.

Nous escomptons toutefois que cette différence de complexité est compensée par la richesse des transitions possibles. En effet, si nous reprenons notre métaphore des jeux de dames et d'échecs, il apparaît que la richesse du jeu d'échecs confère à ce dernier plus d'émules qu'au jeu de dames. De plus, la définition du but n'est pas la même. La stratégie employée aux dames est systématique et consiste à éliminer l'ensemble des pions adverses. La stratégie du joueur d'échecs est plus heuristique, mais peut être réalisée plus rapidement (le « coup du berger » permet de gagner une partie en quatre coups).

De la même manière, le parcours de taxinomies oblige l'individu à « suivre les liens » hiérarchiques jusqu'au but suivant de sa recherche. Le parcours des relations entre les diagrammes de connaissances, par définition plus riches, permet plusieurs types de cheminements pour arriver aux différents diagrammes¹³⁴. Les nombres d'inférences à réaliser peuvent être moins importants que pour le parcours de taxinomies si les liens proposés sont plus « directs » (i.e. plus pertinents pour la tâche de recherche d'information). Plusieurs approches sont envisageables pour cela.

¹³³ Un ou deux types de déplacements sont possibles aux dames. Dans sa thèse, James Allan identifie sept types de liens hypermédias génériques, qu'il qualifie de « pattern-matching links » ([Allan 95], p. 23).

¹³⁴ Pour ces deux types d'outils, la stratégie classique revient en général à passer en revue l'ensemble des éléments disponibles pour sélectionner le plus pertinent. Dans ce cas, l'individu effectue mentalement les inférences correspondantes et évalue la plus pertinente. Une autre stratégie possible est le parcours d'arbre (encore appelé « essai-erreur ») dans laquelle l'utilisateur visualise les destinations des liens et évalue ensuite la pertinence de la cible par rapport à son but. Cette stratégie est souvent qualifiée de « paresse cognitive ».

La première est de présenter en priorité à l'utilisateur les diagrammes utiles de la modélisation (i.e. les étapes informationnelles de l'activité).

Sur ce point, l'utilisation des actinomies peut être intéressante. En effet, les modèles dynamiques de la modélisation offrent des relations d'inférence horizontales plus riches que celles des taxinomies. L'objectif est alors de privilégier et de faciliter la production de ces inférences par l'utilisateur. La première solution envisagée est alors de présenter à l'utilisateur directement les diagrammes correspondant au sous-domaine d'activité dans lequel il a le meilleur niveau d'expertise. L'idée est de privilégier la navigation à l'intérieur de ce sous-domaine. C'est ce qui est réalisé dans le prototype, où les liens inter-sous-domaines sont explicitement mentionnés.

Une autre solution est d'affecter un booléen « utile pour la RI » à chaque diagramme ou à chaque élément de diagramme et de proposer uniquement les diagrammes comportant des éléments utiles. Cette technique n'est intéressante que si elle est mise en œuvre spécifiquement pour chaque type d'utilisateur (i.e. chaque sous-domaine) voire chaque utilisateur. Cela oblige d'une part à rendre invisibles des pans entiers de la modélisation qui peuvent être pertinents pour la recherche d'information dans des cas particuliers. D'autre part, cette technique, pour être efficace, doit mémoriser le comportement et la navigation des utilisateurs du LMCE. Compte tenu des domaines modélisés, nous préférons ne pas avoir recours à des technologies employant des « agents » qui recueillent des informations sur le comportement des utilisateurs¹³⁵. Par ailleurs, l'un des problèmes techniques majeurs posés par cette solution est la conservation de la cohérence de la modélisation pour tous les utilisateurs identifiés. Compte tenu du choix imposé par la méthode MKSM d'employer des liens statiques pour le parcours des relations entre les diagrammes, cette cohérence nous paraît difficile à garantir.

Il est également envisageable de définir des formalismes MKSM spécifiquement adaptés à la navigation cumulative. La difficulté est de mettre au point ces formalismes permettant de conserver la richesse de la modélisation et de l'intégrer dans le processus de recherche

¹³⁵ Le nucléaire est en effet un domaine sensible où toute utilisation à mauvais escient d'une information recueillie sur l'activité des acteurs d'une organisation peut avoir des conséquences fâcheuses. C'est la raison pour laquelle aucune donnée sur cette activité n'est produite sans le contrôle des acteurs de la hiérarchie de cette organisation. C'est notamment la raison pour laquelle le fonds documentaire du SEC est figé au format pdf, les diagrammes de connaissances sont stockés sous forme d'images et les profils d'utilisateurs ne contiennent pas de donnée sur le contenu de l'activité.

d'information¹³⁶. Il ne nous paraît pas évident de satisfaire à ces deux conditions. La solution est alors de réaliser une modélisation des connaissances « à deux vitesses », dont une est adaptée à la formation et l'autre à la stratégie d'accès à l'information cumulative. Compte tenu du surcoût engendré par cette deuxième modélisation, cette solution ne nous paraît pas envisageable.

Une autre approche est de compléter la modélisation existante par une « surcouche » spécifiquement destinée à faciliter l'accès à l'information. Cela permet l'utilisation de cette modélisation dans le double objectif de transmettre le savoir-faire présenté dans les diagrammes et de faciliter le parcours pour la sélection d'éléments pertinents d'un problème d'information. Cette approche nous paraît la plus prometteuse. La condition préalable à sa mise en œuvre est l'identification des diagrammes à considérer comme pertinents, qui est détaillée aux paragraphes 5.1.2.3 et 5.1.2.4. Nous admettons ici pour simplifier que la pertinence d'un diagramme pour ce but est proportionnelle au nombre d'éléments correspondant à des critères d'indexation qu'il contient. Plusieurs solutions techniques sont envisageables, qui ne sont a priori ni exclusives, ni exhaustives.

De tels raccourcis sont implicitement proposés par les cartes conceptuelles spécifiques à la méthode MKSM. Le recours à ces cartes conceptuelles implique cependant la connaissance générique des formalismes MKSM, sinon du contenu des diagrammes. Il est par exemple utile de savoir que les modèles de concepts donnent accès à des critères de fonds, tandis que celui des activités à des critères de forme pour trouver ces types d'éléments. D'où la nécessité d'une formation à l'utilisation du LMCE¹³⁷.

Cette approche peut également être mise en œuvre grâce à des liens sémantiques « raccourcissant » le trajet entre les diagrammes comportant des éléments pertinents pour la recherche d'information. Leur typage spécifique serait alors indiqué dans le multiplexeur de liens, comme pour les autres liens de navigation dans la collection de diagrammes. Cette solution présente l'avantage d'être facile à mettre en œuvre a posteriori.

La particularisation des éléments des diagrammes porteurs de liens vers les diagrammes pertinents pour la recherche d'information mérite également d'être étudiée. Cette solution permettrait de « flécher » le plus court chemin passant par les diagrammes les plus susceptibles de contenir des éléments du besoin d'information de chaque type d'utilisateur.

¹³⁶ Nous avons vu qu'au chapitre 2 que la mise au point de tels formalismes est complexe et est du ressort de la diagrammatique. Elle oblige de plus à redéfinir la systématique des relations existant entre les différents formalismes de la méthode MKSM.

¹³⁷ C'est d'ailleurs dans cette optique que les noms choisis pour les différents éléments de l'interface (caddie à concepts, multiplexeur de liens) sont volontairement « parlants » pour les utilisateurs.

En l'état actuel du processus de génération des diagrammes à l'aide du logiciel Visio®, cette particularisation doit avoir lieu en amont de la génération de la collection de diagrammes sous forme d'images. Il peut par exemple s'agir de modifier la couleur ou l'ombrage des éléments à la source des liens à particulariser. Cette solution pose le problème de la validation de la nouvelle collection de diagrammes. Dans le cadre du projet-pilote présenté ici, nous ne disposons plus des filtres nécessaires à cette « régénération » des diagrammes.

Pour conclure ce paragraphe, nous constatons que la gestion des connaissances n'est a priori pas destinée à permettre l'accès à l'information. Toutefois, la modélisation résultant d'un projet de gestion des connaissances comporte tous les éléments susceptibles de permettre cet accès à l'information. L'organisation de ces éléments dans la collection de diagrammes les rend conceptuellement plus complexes à manipuler. D'où l'intérêt de l'ergonomie pour aider l'utilisateur à maîtriser cette complexité. Cela peut être réalisé en présentant en priorité les diagrammes supposés pertinents pour l'accès à l'information et en facilitant cette manipulation à l'aide d'outils spécifiquement destinés à cette tâche. De tels outils sont les cartes conceptuelles, les « raccourcis » sémantiques vers les diagrammes les plus riches en termes d'éléments significatifs pour la composition d'une requête informationnelle ou le « fléchage » de ces diagrammes. Le premier de ces types d'outils fait déjà partie du prototype. La mise en place du deuxième nous paraît envisageable à court terme. Là encore, une évaluation du prototype spécifié au chapitre 4 nous éclairerait sur l'efficacité de ces solutions.

5.1.3. *Apport de l'indexation conceptuelle pour la formalisation du besoin d'information*

La recherche d'information dans le LMCE fait appel à un moteur de recherche en texte intégral. Moyennant une reformulation appropriée, tous les contenus textuels des éléments des diagrammes sont donc supposés atteignables par cet outil. Dans ce paragraphe, nous justifions en termes de pertinence pour l'accès à l'information, le choix fait de recourir à des index pour accéder aux documents. Deux éléments de réponse sont à considérer.

D'une part, l'indexation séparée des documents en fonction de leur appartenance aux critères identifiés permet le filtrage des documents et diminue le bruit (i.e. le nombre de documents non pertinents retournés). Elle permet par exemple d'éliminer les documents comportant les termes de la requête dans des références ou des fragments non significatifs pour

l'indexation¹³⁸. En ce sens, la séparation des index se substitue à la segmentation documentaire telle qu'elle est envisagée dans [Lainé-Cruzel & al. 96] ou au « cluster-based search » de [Salton & Mc Gill 83] (via [Iwayama & Tokunaga 95], p. 273). Cette technique nous permet de contourner l'indivisibilité des documents inhérente à leur présentation au format pdf.

D'autre part, dans le projet-pilote, l'indexation conceptuelle des documents présentée ici correspond à un classement indiqué par les experts de l'unité modélisée. Si nous reprenons la définition de l'indexation conceptuelle selon Yannick Prié ([Prié 01]), celle-ci doit permettre des inférences sur les index que ne permet pas une recherche plein texte classique¹³⁹. Ces inférences sont supposées réalisables autant par l'outil d'accès à l'information que par son utilisateur. L'indexation conceptuelle permet alors de faire correspondre la vision du système d'information de l'expert avec son utilisation par le système. La principale inférence de ce type est la décomposition structurelle de ces index dans le modèle de l'information présenté. Par exemple, la structure arborescente des index permet au système d'éviter la recherche dans les sous-index. De même, nous avons vu au chapitre 4 que l'intersection des index doit être traitée sous forme d'inférences spécifiques par le système de règles.

Nous pensons que ces deux raisons justifient l'emploi d'index, malgré le surcoût qu'implique leur réalisation en terme de charge de travail pour le chargé de projet et pour les experts modélisés. Cette représentation du système d'information permet un accès à l'information plus fiable que la recherche en texte intégral sur les différentes valeurs de reformulation des index. D'autre part, elle est plus facilement opérationnalisable, voire perfectible, par le système.

5.1.4. *Apports spécifiques à la gestion des connaissances*

Une partie importante de ce travail de thèse a été réalisé à partir de la méthode de gestion des connaissances MKSM développée au CEA. Parallèlement à sa problématique d'accès à l'information, cette thèse s'inscrit également dans le cadre de travaux de recherche

¹³⁸ Par exemple, l'interrogation dans la base pdf du SEC de la totalité du fonds documentaire sur un type quelconque de matière fissile retourne nécessairement le standard de criticité (ce document donne la liste des valeurs minimales critiques pour chaque configuration connue de la matière fissile). Inversement, la plupart des notes techniques du SEC font référence à ce document. Une recherche sur son titre retourne donc une précision peu performante.

¹³⁹ Des travaux plus poussés ont cependant été réalisés. Dans sa thèse, Jean-Luc Simoni ([Simoni 00]) ajoute par exemple à la recherche plein texte de SPIRIT® une surcouche permettant la construction de graphes des termes significatifs des documents pour permettre justement ces inférences.

spécifiques à cette méthode. Plusieurs questions dans ce domaine ont été soulevées durant cette thèse, auxquelles nous proposons des éléments de réponse dans cette partie.

Elle s'articule en trois sous-parties. Nous abordons tout d'abord le positionnement de la méthode MKSM dans un cadre plus général. Nous discutons les paradigmes systémique et sémiotique sur lesquels cette méthode est bâtie. Cette discussion nous permet d'aboutir à une représentation plus pragmatique du positionnement des modèles. Celle-ci peut être utilisée pour présenter la structure hypermédia du modèle du domaine à l'utilisateur du système d'accès à l'information réalisé. Dans une deuxième sous-partie, la pertinence de l'utilisation de la méthode de gestion des connaissances MKSM pour la démarche transdisciplinaire de conception de l'outil d'accès à l'information est évaluée. Il s'agit de définir l'utilisabilité de chacun des formalismes diagrammatiques pour les différentes étapes de la mise au point des modèles conceptuels et de l'interface. Dans un troisième temps, des éléments de réflexion plus théoriques sont proposés afin d'évaluer la pertinence de l'approche récursive qui consiste à réinjecter dans l'activité d'une organisation la modélisation de cette activité obtenue en gestion des connaissances.

5.1.4.1. Éléments de réflexion sur la présentation et le positionnement de la méthode de gestion des connaissances MKSM

Dans cette partie, nous revenons à la définition de la méthode MKSM afin d'apporter quelques considérations sur les fondements de cette méthode. Ces considérations découlent de nos réflexions ainsi que des enseignements tirés du travail présenté dans ce manuscrit. Il s'agit de discuter le statut des modèles MKSM et des paradigmes systémique et sémiotique sur lesquels cette méthode se fonde par rapport au microscope dans l'optique d'une présentation ergonomique du référentiel d'accès à l'information avec le LMCE.

La méthode MKSM définie par [Ermine & al. 96] propose différents formalismes diagrammatiques. Elle les organise suivant une double décomposition systémique et sémiotique du « système de connaissances » de l'organisation modélisée. Cette décomposition est représentée par un outil conceptuel : le microscope de la connaissance. Nous avons vu au chapitre 3 que ce concept de système de connaissances est notamment utilisé pour déterminer, en présence des instances de décision d'une organisation, l'opportunité d'y mener un projet de gestion des connaissances.

Il se situe donc à un niveau d'abstraction qui est sans doute nécessaire pour se faire une idée a priori de l'objet d'une telle modélisation. Néanmoins, d'un point de vue opérationnel, la

notion de système de connaissances est trop immatérielle pour servir de base à la modélisation. En particulier, le discours tenu aux experts lors de la phase d'interviews n'est pas cohérent avec cette notion. Le chargé de projet se place alors sur un plan purement informationnel et parle d'expertise, d'information et de rédaction technique.

Notre objectif dans ce paragraphe est de proposer et de justifier une réorganisation de ces modèles, adaptée à notre approche informationnelle et au caractère pluri-individuel de l'activité décrite.

Dans les projets antérieurs, la modélisation des connaissances se traduisait factuellement par une modélisation de l'activité d'un ou de plusieurs experts autour desquels est centrée l'activité de l'organisation. C'est pourquoi il nous semble plus exact de redéfinir l'origine du référentiel de la méthode MKSM (i.e. le macroscopie) en la centrant sur ces experts (ou plus généralement sur les acteurs de l'organisation) plutôt que sur leurs systèmes de connaissances.

Nous avons en outre souligné que la seule trace utilisable du savoir-faire des individus dans l'organisation (et donc de leurs « connaissances » dans le domaine modélisé) est l'information circulant lors des processus de communication impliquant cet individu. C'est donc cette information (dont la décomposition systémique est sa perception du système d'information) qui doit être considérée comme le « signe » dont on cherche à établir une description sémiotique. Par ailleurs, ce recentrage du macroscopie se justifie également par le fait qu'il s'agit dans cette thèse d'utiliser la modélisation d'une activité pour faciliter l'accès à l'information dans l'exercice de cette activité.

Cette redéfinition du macroscopie nous amène donc à repositionner les modèles les uns par rapport aux autres. Ce repositionnement est proposé dans la figure ci-dessous.

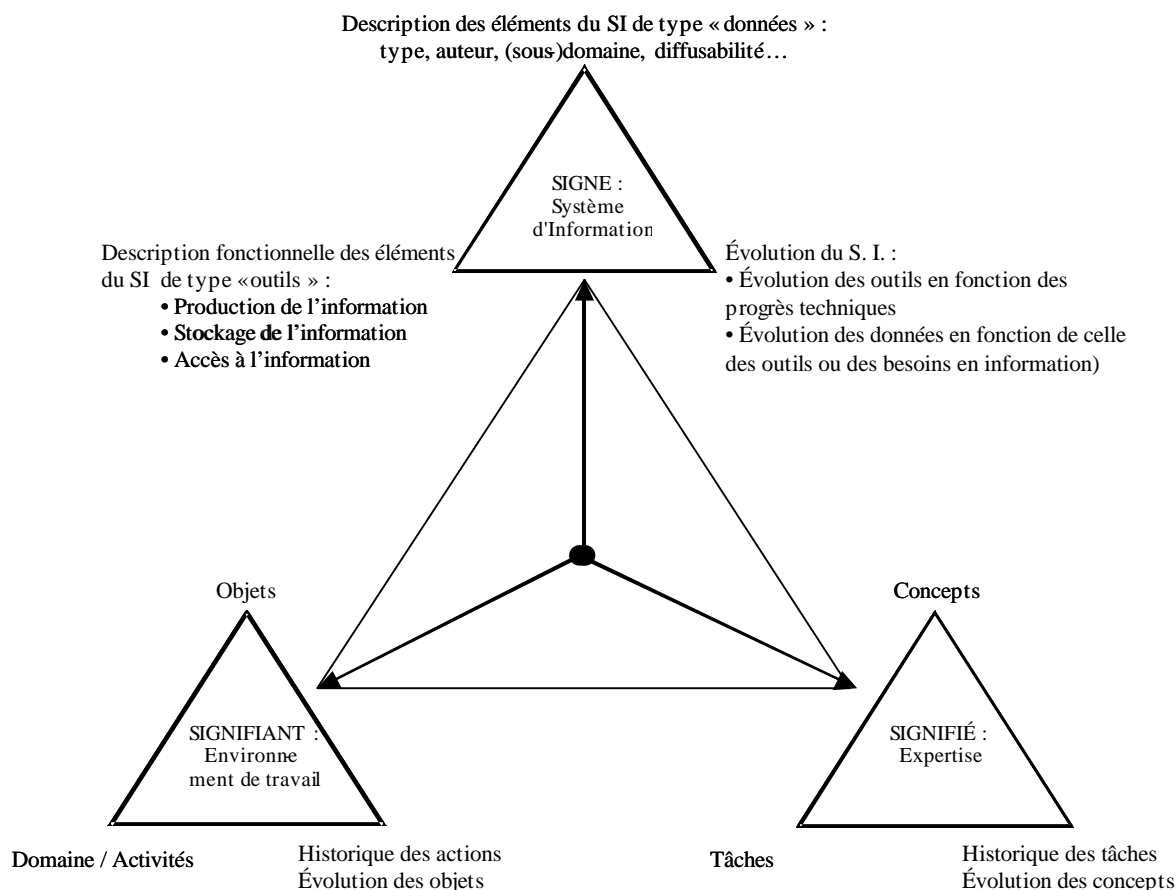


Figure 55. Une tentative de redéfinition du macroscopie MKSM.

En comparant cette version du macroscopie avec celle de [Ermine & al. 96] présentée au chapitre 3, nous voyons que peu de choses ont changé. Les modifications apportées aux trois points de vue systémiques liés au système d'information sont spécifiques à notre problématique. Elles dénotent donc la subjectivité de cet outil en fonction du but de la modélisation.

Nous avons également ajouté des formalismes dynamiques diachroniques correspondant aux points de vue systémiques structurels et synchroniques modélisés. Deux de ces formalismes (historique des actions et évolution des concepts) sont ceux présentés par [Van Craeynest & al. 00]. Nous avons également fait figurer sur ce diagramme des formalismes qui n'existent pas (i.e. évolution des objets et historique des tâches). L'« invention » de ces formalismes se justifie par une tentative de symétrisation des points de vue diachroniques du macroscopie liés aux points de vue sémiotiques signifiant et signifié. En effet, sans ces formalismes, un « trou » apparaît dans la structure de cet outil conceptuel.

Nous avons en outre fait figurer le modèle des objets, comme point de vue de description structurelle de l'environnement de travail de l'individu. L'adjonction de ce modèle en lieu et

place de celui du domaine nous conduit à nous interroger sur la place que ce dernier doit alors prendre dans le macroscopie. Nous avons déjà souligné dans le chapitre 3 que la présence d'un modèle dynamique ne se conçoit que dans un sommet fonctionnel de l'un des triangles systémiques du macroscopie. De plus, ce modèle ne concerne pas l'expertise modélisée, mais le contexte de travail de cette expertise. C'est pourquoi nous lui attribuons la même place que le modèle des activités.

Toutefois, le problème du positionnement du modèle du domaine n'est pas résolu pour autant. En effet, nous voyons que si nous supprimons ce modèle du macroscopie, nous obtenons une décomposition symétrique du contexte de travail et de l'expertise des individus dans l'organisation. Une telle décomposition positionne alors le macroscopie comme un outil conceptuel permettant aux utilisateurs de visualiser le passage entre les deux points de vue sémiotiques. Dès lors, cet outil peut être utilisé comme support d'explicitation de la fonction de ces modèles pour le lecteur du livre de connaissances. De plus, cet outil devient indépendant du niveau de granularité de la modélisation : il est valable au niveau du domaine modélisé comme aux niveaux de chacun des sous-domaines. Les points de vue informationnels considérés sont alors la totalité ou les parties correspondant aux différents sous-domaines du système d'information de l'organisation. La figure ci-dessous indique alors une présentation plus utilisable de cet outil.

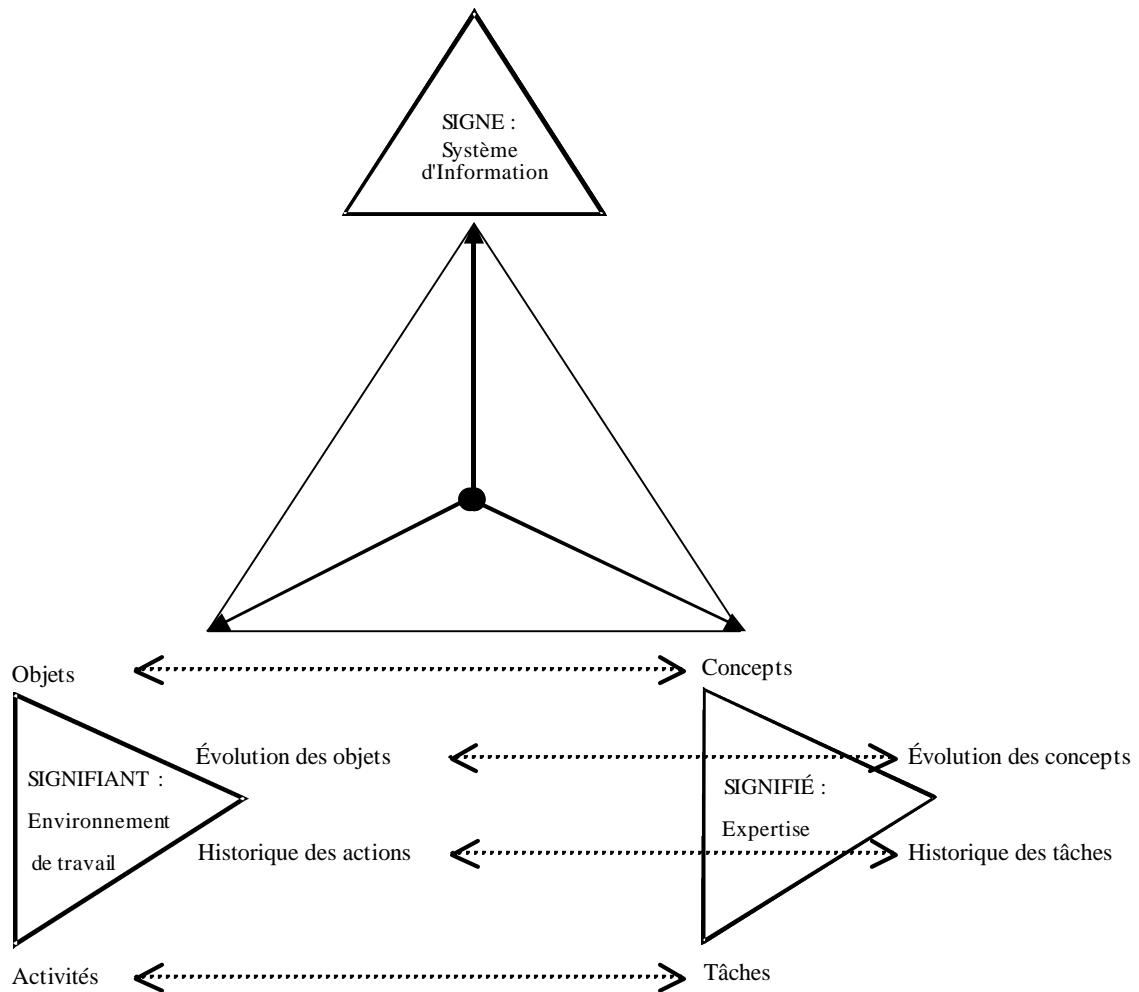


Figure 56. Une deuxième tentative de redéfinition du macroscopie prenant en compte l'utilisabilité de cet outil pour sa présentation dans le LMCE.

Sur cette figure, nous avons fait pivoter les triangles systémiques pour des raisons de lisibilité¹⁴⁰. Le fait qu'il existe désormais deux formalismes diachroniques pour chaque décomposition systémique nous conduit cependant à nous interroger sur la pertinence du triangle comme moyen de représentation de cette dimension systémique. Plus précisément, il nous semble que celle-ci serait plus claire si les distinctions structurelle / fonctionnelle et synchronique / diachronique étaient placées au même niveau. Nous proposons donc la représentation suivante des points de vue systémiques de la méthode MKSM :

¹⁴⁰ Il est intéressant de remarquer que cette opération « déstabilise graphiquement » tout l'édifice du macroscopie. Jean-Marie Van Craeynest, auteur des formalismes diachroniques, fait d'ailleurs remarquer qu'une telle opération le rend « invendable » auprès des responsables des unités où sont susceptibles d'être menés des projets.

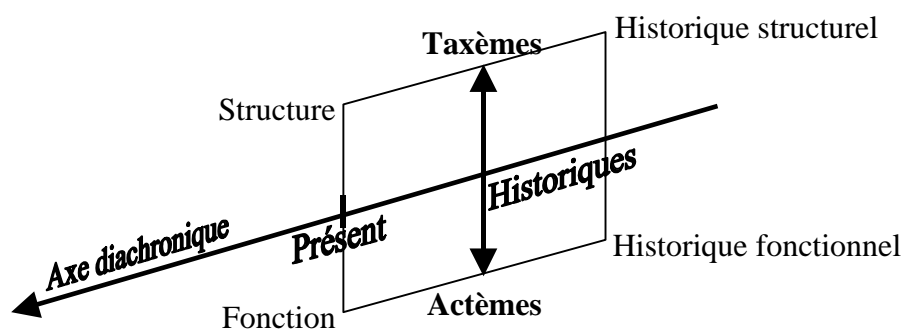


Figure 57. Proposition de décomposition systémique bidimensionnelle.

Sur cette figure apparaissent deux axes, l'un « quantique » et l'autre continu. La distinction qui est faite entre les points de vue classificatoires et dynamiques est dichotomique, tandis que le positionnement des modèles sur l'axe diachronique (i.e. le temps) est continu. En particulier, si le point de fonctionnement synchronique (présent) du système se situe en un point fixe de cet axe, les points de vue historiques représentent des intervalles (non nécessairement fermés) de cet axe. Ces points de vue sont d'ailleurs supposés décrire l'évolution du système jusqu'au point synchronique pour avoir du sens. En ce sens, la représentation schématique ci-dessus est donc inexacte. Toutefois, nous la conservons telle quelle, en gardant à l'esprit cette inexactitude.

Un domaine d'activité est alors composé de deux de ces parallélogrammes, l'un pour le contexte, l'autre pour l'expertise. Cette distinction est héritée du paradigme sémiotique sur lequel s'appuie la méthode MKSM. Elle est difficile à conserver si elle est décorrélée du signe (i.e. l'information), qui donne sens à cette théorie. Nous préférons alors parler d'une distinction entre les niveaux concret (les objets et la production de flux matériels dans l'activité) et abstrait (i.e. les concepts et les stratégies cognitives des experts). Nous introduisons cette deuxième dichotomie dans la figure suivante.

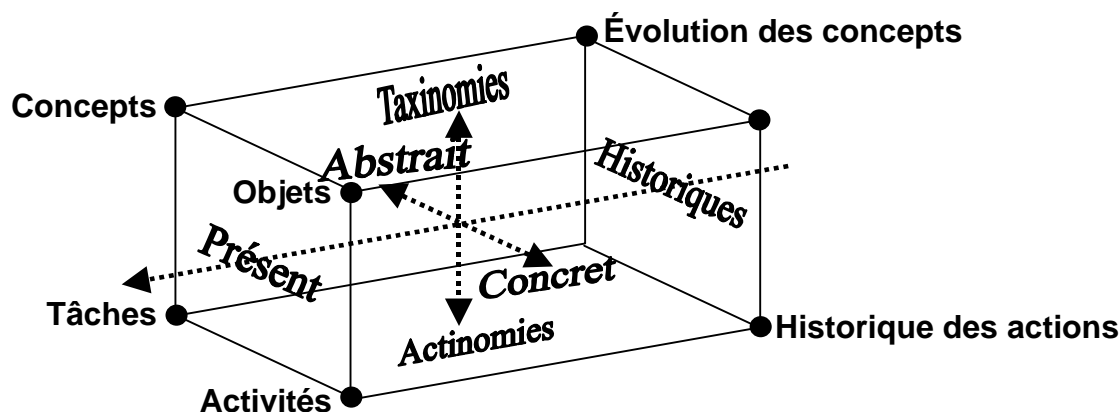


Figure 58. Représentation systémique des points de vue MKSM d'un domaine.

Chacun des sommets de ce cuboï de représente alors un des points de vue MKSM, matérialisé par un point gras. Ce schéma permet alors de représenter chacun des sous-domaines modélisés, de sorte que l'activité liée à la maîtrise du risque de criticité peut être représentée dans le cas du SEC de la façon suivante :

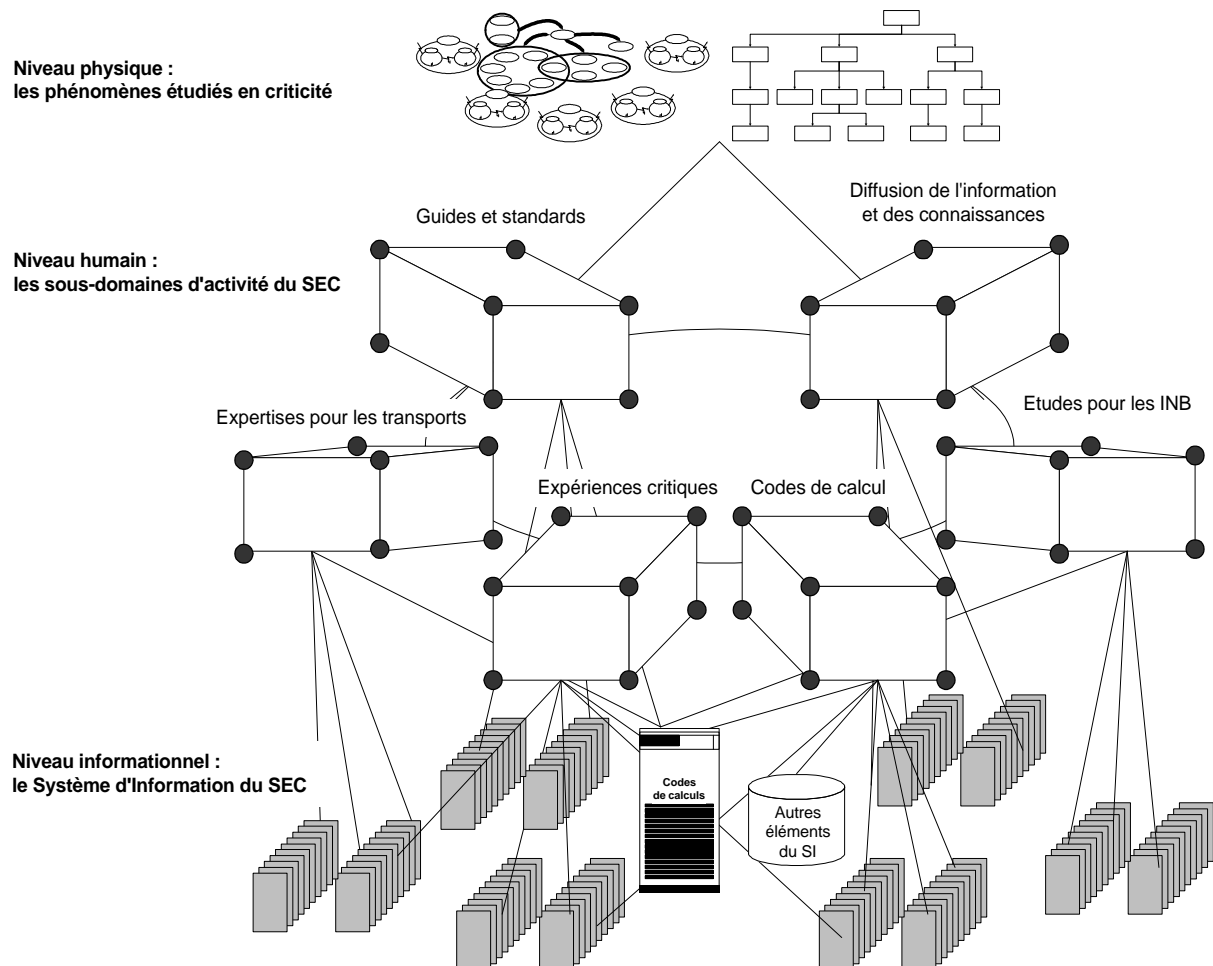


Figure 59. Le macroscopie de l'accès à l'information : une vision à trois niveaux d'une activité multiple autour d'un domaine scientifique et technique.

Ce diagramme repositionne les différents sous-domaines d'une activité scientifique et technique autour d'un phénomène identifié. Les décompositions internes des sous-domaines ne sont pas représentés ici, pour des raisons de lisibilité. Il apparaît que le paradigme sémiotique sur lequel s'appuie la méthode MKSM n'est plus nécessaire ici.

Les trois niveaux identifiés permettent notamment de différencier les diagrammes de niveaux phénoménologiques et organisationnels de l'activité décrite. Nous avons à ce titre placé les modèles des processus et des concepts généraux (dans notre exemple, la catégorisation des matières fissiles) au niveau physique. Les points de vue spécifiques à l'activité sont situés dans les cuboï des représentant les différents sous-domaines.

Le système d'information est le troisième niveau de ce microscope. Il est accédé à partir des différents types de diagrammes selon les liens spécifiques à chaque sous-domaine. Les liens de navigation inter-modèles ne sont pas indiqués sur ce diagramme.

Finalement, cette discussion sur le positionnement de la méthode MKSM permet de nous conforter dans l'utilisation du paradigme systémique pour la description de l'activité des utilisateurs du LMCE. Il apparaît que la présentation du microscope telle qu'elle est présentée dans [Ermine & al. 96] n'est plus satisfaisante dans le cas où plusieurs sous-domaines sont modélisés. Le repositionnement des modèles dans le contexte global du domaine scientifique et technique étudié conduit à l'abandon du paradigme sémiotique. Celui-ci est remplacé par une vision à trois niveaux dont deux correspondent à la distinction entre le domaine étudié et l'activité des individus.

Cette activité est alors positionnée à la fois en fonction du domaine et de l'accès à l'information. D'un point de vue opérationnel, cette version du microscope peut être utilisée pour présenter à l'utilisateur la structure conceptuelle du LMCE. En ce sens, et d'après [Bernstein 90], via [Nanard & Nanard 91], il constitue un moyen d'éviter ou de limiter les phénomènes de désorientation lors de la navigation hypermédia dans cet outil.

5.1.4.2. Discussion du paradigme d'utilisation de la gestion des connaissances comme référentiel de conception et d'utilisation d'un système d'accès à l'information scientifique et technique (Interactions entre les disciplines pour la conception)

Le but de ce paragraphe est d'aborder les liens qui existent entre la gestion des connaissances et les autres composantes de l'accès à l'information avec le LMCE concernant la phase de modélisation. Il s'agit d'identifier les similarités entre les trois disciplines, afin d'essayer d'améliorer le processus de conception du LMCE, et, s'il y a lieu, le LMCE lui-même.

En particulier, les modèles MKSM sont utilisés dans la phase de conception du LMCE pour mettre au point les deux autres modèles conceptuels. Des exemples d'utilisation de ces formalismes dans ce but, tels que l'analyse des tâches des utilisateurs ou la définition du modèle de l'information du SEC sont donnés au chapitre 4. Ils sont justifiés dans cette partie.

Ce paragraphe commence par replacer les éléments de la modélisation dans leur contexte théorique. Il passe ensuite en revue l'intérêt de chacun des formalismes MKSM pour la conception du LMCE.

5.1.4.2.1. *Les aspects théoriques des modélisations*

En gestion des connaissances comme en ergonomie, la phase de modélisation se situe en amont de tout processus de conception. Elle se fonde, entre autres, sur plusieurs domaines de recherche, qu'il est commode de regrouper sous la dénomination de sciences cognitives. Cette appellation réunit des disciplines qui s'appuient sur des bases théoriques et des courants de pensée souvent très divers, voire contradictoires [Ganascia 93]. Notre but ici n'est pas de nous inscrire dans l'un ou l'autre de ces différents courants, mais au contraire d'en identifier les similarités pour mieux évaluer l'ensemble des potentialités des modélisations effectuées.

La méthode de gestion des connaissances MKSM se situe actuellement aux confluent de la systémique et de la sémiotique, tentant à la fois de restituer une vision globale des systèmes de connaissances, et de leur associer un sens en les replaçant dans leur contexte. C'est dans cette optique qu'ont été retenus plusieurs points de vue, pour chacun desquels un modèle diagrammatique adéquat a été adopté.

Ainsi modélisée, notre description du système de connaissances peut être utilisée pour spécifier l'outil LMCE. Ces spécifications comportent d'une part les fonctionnalités du LMCE : visualisation et parcours de la collection de diagrammes et liens avec les outils de l'organisation utilisés pour les principales tâches modélisées. D'autre part, elles permettent de définir les interfaces appropriées (à partir des tâches effectives modélisées de l'utilisateur, d'une vision systémique globale du sens qu'il associe au système d'information et de l'environnement dans lequel il évolue).

Chacun de ces modèles permet donc de décrire un aspect de l'activité modélisée : d'un point de vue opérationnel¹⁴¹, il possède une *fonction de caractérisation* d'un élément du système de connaissances. Pour identifier chacune de ces fonctions, nous nous fondons sur les formalismes employés dans la première version de la méthode MKSM, décrite dans [Ermine & al. 96]. Nous avons en effet choisi de ne pas traiter ici les formalismes d'évolution des concepts et d'historique des actions (présentés dans [Van Craeynest & al. 00]), car ces formalismes dynamiques diachroniques ne nous sont pas apparus comme très utiles pour la spécification du LMCE – même s'ils présentent une réelle avancée pour la compréhension par le lecteur de la collection de diagrammes. En effet, ces formalismes, sont destinés à tracer

¹⁴¹ Le but de ce paragraphe n'est cependant pas ici d'opérationnaliser la méthode MKSM. [Charreton 98] a montré les avantages et les limites de cette approche. Nous pensons quant à nous, que la nature qualitative des formalismes MKSM ne se prête pas à une automatisation des tâches accomplies par les experts interviewés ni ne permet le remplacement de ceux-ci. Notre idée est que ces formalismes peuvent être interprétés par le concepteur du LMCE pour en tirer les informations nécessaires à tout processus de conception logicielle.

l'évolution au cours du temps de notions par ailleurs décrites dans d'autres modèles sous leur forme actuelle. Ces modèles pourraient cependant ultérieurement être utilisés pour tracer l'évolution du LMCE (« versionning ») comparativement à celle du système de connaissances qu'il décrit.

Dans cette partie, nous passons en revue l'ensemble des points de vue modélisés. L'objectif est de proposer une méthodologie d'interprétation permettant, à partir de ces modèles, d'en extraire une fonction générique au sein du LMCE.

5.1.4.2.2. *Du modèle des activités à la spécification des fonctionnalités du LMCE*

Le modèle d'activités se fonde sur le formalisme SADT qui présente une vision fonctionnelle d'un système modélisé. Il permet notamment d'identifier les buts globaux des acteurs de l'organisation dans l'accomplissement de leur activité. De plus, les flux documentaires circulant entre les différentes activités apparaissent dans ce modèle. Ils permettent de visualiser le type de livrable ciblé par chacun de ces buts. La connaissance de ces buts permet la spécification des objectifs génériques d'un outil susceptible d'aider ces acteurs dans l'accomplissement de leur activité.

La structure arborescente de ce formalisme permet alors de choisir le niveau de granularité auquel l'objectif du système doit être défini. Cela permet de passer de la définition d'un objectif à celle des fonctionnalités nécessaires pour le réaliser. Dès lors, le formalisme des activités peut être utilisé pour décrire non plus l'activité de l'organisation modélisée, mais celle d'un acteur utilisant le système réalisé pour l'accomplissement de cette activité. En d'autres termes, le formalisme des activités permet de présenter les spécifications des fonctionnalités du système, comme cela est fait au chapitre précédent.

Nous pouvons ici noter que ces spécifications sont indépendantes du type d'objectif identifié pour le système puisque le formalisme SADT est à l'origine une méthode de spécification en génie logiciel. Par conséquent, l'accès à l'information est une application parmi tant d'autres de cette méthode. Cependant, cette méthode présente un point de vue fonctionnel qui ne permet pas la description des stratégies cognitives des individus dans l'accomplissement de leur activité. Cette description est donnée par la modélisation des tâches, selon un formalisme MKSM issu de ceux utilisés en ergonomie.

5.1.4.2.3. *De la modélisation des tâches à la définition de l'interface utilisateur*

En ergonomie, la recherche s'intéresse principalement aux aspects cognitifs du comportement d'un utilisateur devant un artefact (dans notre cas, un logiciel). Cette approche a permis de définir des méthodes de modélisation des tâches cognitives, utilisées pour la conception d'IHM. En particulier, [Tricot & Nanard 98] étudient l'exemple de deux de ces méthodes (GOMS et MAD) dans le cas de la conception d'hypermédias de recherche d'information. Les conclusions que tirent ces auteurs sont prudentes mais penchent néanmoins en faveur de l'utilisation de MAD pour la conception de ce type d'outils. Ce formalisme présente l'avantage de donner une description complète des tâches, incluant entre autres leurs états initiaux et finaux ainsi que les données utilisées pour leur réalisation.

Ce formalisme est à la base du modèle des tâches MKSM présenté dans [Ermine & al. 96]. Dans sa version originelle, le modèle des tâches MKSM reprend d'ailleurs une description diagrammatique de l'ensemble de ces paramètres. Ceux-ci ont depuis été supprimés car ils étaient en doublon avec les flots de données mentionnés dans le formalisme des activités et nuisaient à la lisibilité des diagrammes. Par conséquent, le formalisme des tâches MKSM, associé à celui des activités, permet la présentation des tâches cognitives des utilisateurs pour la conception d'un système.

Les détracteurs de ce type de méthode objectent cependant que la modélisation des tâches ne tient pas compte de certaines contraintes, extérieures à l'environnement logiciel, auxquelles est soumis l'utilisateur durant son interaction avec le système. Ces contraintes peuvent, pour certains types d'applications, modifier voire déterminer les modèles d'utilisation. C'est pourquoi le fait de replacer l'activité d'utilisation du logiciel dans son contexte¹⁴² - quand celui-ci est identifié - peut permettre une meilleure intégration du logiciel dans l'environnement de travail et donc améliorer l'interaction de l'utilisateur avec le programme. Cette approche, sans doute un peu téméraire, se justifie cependant plus particulièrement dans le cas où le logiciel à réaliser est un « méta-programme », permettant d'accéder à d'autres logiciels, ayant chacun leur propre interface¹⁴³. Qui plus est, chaque type d'utilisateur, s'il est identifié, accomplit non seulement des tâches différentes, mais se construit aussi ses propres

¹⁴² C'est-à-dire, pour un programme à usage professionnel, l'environnement de travail, mais aussi le système d'exploitation de l'ordinateur, et les autres applications susceptibles d'être utilisées en même temps.

¹⁴³ Nous préférons ici la notion de méta-programme à celle de système d'exploitation, car du point de vue de l'ensemble des logiciels proposés aux acteurs de l'organisation, le LMCE se situe au même niveau que les autres logiciels. Il se contente de proposer des liens vers ceux-ci et n'en contrôle pas l'exécution.

schémas cognitifs, correspondant à sa propre activité dans un contexte donné (qui peut, lui, être commun à plusieurs types d'utilisateurs).

Par conséquent, il semble intéressant d'étudier la possibilité de mener en parallèle les modélisations en gestion des connaissances et en ergonomie. En effet, le Livre de Connaissances Électronique doit être conçu entre autres pour faciliter l'accès à l'information aux différents experts. Ces experts, dont il faut capitaliser les connaissances, sont donc également des destinataires du logiciel, et constituent à ce titre, une classe d'utilisation, pour laquelle il s'agit de proposer une interface spécifique, correspondant à la fois à leur activité, et harmonisée avec celles des logiciels qu'ils utilisent. Dans la mesure où l'utilisation des outils informatiques fait partie de l'expertise du domaine, la modélisation de cette activité doit être incluse dans la modélisation en gestion des connaissances. Il n'est alors pas surprenant de vouloir la réutiliser pour extraire les « profils cognitifs » de ces experts à des fins ergonomiques.

Pour faciliter cette mise en correspondance des deux disciplines, il est alors nécessaire d'« unifier » les bases théoriques des modélisations. Pour cela, nous rappelons que la méthode MKSM s'appuie sur la sémiotique, qui associe au signe (l'information), un signifiant (le contexte dans lequel elle a été émise), et un signifié (le sens qui lui est associé). L'expérience prouve que cette approche, associée à la systémique, facilite la restitution de la globalité des interactions existant à l'intérieur d'un système de connaissances. Cependant, dans les cas où plusieurs expertises sont modélisées, l'organisation peut être considérée comme un « système d'acteurs », évoluant dans un même contexte, mais associant à une information un sens différent. Les systèmes d'informations (entre acteurs ou entre un acteur et l'extérieur du système), sont désormais particuliers aux acteurs, et définissent une partie (explicitée ou quantifiable) des flux de connaissances circulant dans l'organisation.

Cette vision du système de connaissances est dès lors beaucoup plus facilement adaptable à l'ergonomie, dans la mesure où chaque acteur appartient à une classe d'utilisateurs, avec un niveau d'expertise plus ou moins élevé. Le contexte, lui, représente une modélisation du système, à un niveau de granularité plus grossier, c'est-à-dire le point de vue de l'observateur extérieur, non-initié à l'activité commune qui réunit les acteurs. La modélisation du contexte correspond donc également à une autre classe d'utilisation identifiée, même si elle ne définit qu'un type d'utilisation du Livre de connaissances électronique, plus général que ceux définis pour les acteurs. Il convient de recueillir des informations sur les objectifs d'utilisation du Livre de connaissances électronique par des utilisateurs extérieurs à l'organisation et non

clairement identifiés. Ce type d'informations ne peut par définition pas être trouvé dans la modélisation.

5.1.4.2.4. Du modèle des concepts à l'indexation sémantique de l'information

Le modèle des concepts est un modèle arborescent statique qui décrit la façon dont un expert du domaine classifie un système de notions. L'ensemble des diagrammes de concepts issus d'un projet de gestion des connaissances constitue sa taxinomie du domaine (Voir plus haut). Nous avons vu au chapitre 2 que ce savoir classificatoire est couramment utilisé pour fournir des ontologies du domaine pour plusieurs types d'indexation documentaire : structurée, à rôles ou conceptuelle.

Dans les différents travaux sur l'indexation documentaire que nous avons trouvés dans la littérature, les représentations les plus évoluées des corpus documentaires s'appuient sur des graphes arborescents, comme l'indexation structurée (cf. [Roussey & al. 99]). C'est pourquoi nous n'avons pas tenu compte des autres types de graphes utilisés dans les modèles MKSM (et notamment les modèles dynamiques) pour l'indexation sémantique¹⁴⁴ du système d'information. En effet, nous voyons mal comment répercuter pour l'indexation des relations dynamiques telles que le déclenchement d'un processus ou l'utilisation d'une ressource. Cet aspect de l'indexation conceptuelle est sans doute à approfondir, ce qui n'est pas fait dans ce travail¹⁴⁵. Celui-ci se limite à définir une approche de la mise en relation de la méthode MKSM et d'un système d'information en utilisant une décomposition arborescente comme représentation pivot du domaine.

Le modèle des concepts présente une telle décomposition des notions utilisées dans le fonds documentaire à indexer. Cette classification sémantique permet de faire le lien entre les diagrammes de connaissances et le fonds documentaire. Un modèle des concepts constitue donc une carte conceptuelle orientée-métier du système d'information. Deux questions concernant ces modèles sont alors à régler.

¹⁴⁴ S'ils ne sont pas utilisés pour la construction des index, les textes contenus dans les éléments des diagrammes peuvent cependant entrer dans la construction des requêtes documentaires. En outre, chacun de ces éléments possède un type qui est utilisé dans la sélection des critères de forme de l'information recherchée.

¹⁴⁵ La richesse de l'ensemble des relations proposées dans la méthode MKSM apparaît toutefois en aval de l'indexation documentaire, lors de la définition du besoin d'information grâce à la stratégie cumulative. Cela permet entre autres de limiter la complexité de la recherche d'information dans le cas de l'utilisation des autres stratégies d'accès à l'information disponibles avec le LMCE.

Un premier problème pratique qui se pose est la façon de traiter les multiples décompositions structurelles d'une même notion du domaine. Cela est illustré par l'exemple des INB (installations nucléaires de base) décrites dans notre projet-pilote à la fois en tant que concepts et qu'objets physiques. Cette distinction est introduite par la création d'un modèle des objets, similaire au modèle des concepts dans sa forme et destiné à représenter ce type de contenu. Dans ce cas, nous avons choisi de privilégier l'indexation en fonction du modèle conceptuel de la notion, puisque c'est celle qui est censée être directement issue de la représentation que se font les experts de cette notion.

Le deuxième problème est de définir la priorité relative des différentes notions du domaine. La plupart des domaines modélisés dans les projets de gestion des connaissances nécessitent la description conceptuelle de plusieurs notions indépendantes. Les arbres conceptuels résultants sont a priori des taxinomies en fonction desquelles une indexation conceptuelle du fonds documentaire du sous-domaine est pertinente. Un document, en tant qu'objet, hérite alors des propriétés des différentes décompositions existantes. Or, la notion d'« héritage multiple » n'existe pas dans les formalismes arborescents de réseaux sémantiques dont est issu le modèle des concepts MKSM. La mise en œuvre de cette représentation sous forme d'une arborescence de fichiers d'index ne se prête pas non plus à une telle définition. Il faut donc « fusionner » ces arbres pour obtenir une décomposition d'un document en fonction de toutes les notions pertinentes du domaine.

L'arbre conceptuel résultant de la fusion de ces taxinomies a alors la complexité (i.e. le nombre de niveaux de décomposition) des arbres additionnés et un nombre d'éléments égal au produit des nombres d'éléments de chacun des arbres. Compte tenu du nombre variable mais a priori non limité de notions à indexer dans un projet de gestion des connaissances, il convient de limiter le nombre de critères d'indexation. L'objectif est d'éviter une décomposition – et par conséquent un nombre d'index – trop important. Cela rendrait l'opération de sélection des index lors de la formalisation de la requête complexe, la recherche d'information dans ces index inutilement longue et la mise à jour de ces index fastidieuse. Dans ces conditions, nous avons choisi de privilégier l'une de ces notions pour l'indexation conceptuelle, les autres étant traitées par la recherche plein texte. Différents critères pour le choix de cette notion sont envisageables. La complexité des arbres ou la quantité de bruit généré par la recherche documentaire sur les éléments de ces arbres dans la totalité du fonds sont des éléments facilement mesurables. D'un point de vue ergonomique, la phase de modélisation permet d'identifier la notion dont les éléments sont les plus fréquemment

employés pour la recherche documentaire ou la notion la plus importante aux yeux des experts. Si une classification des documents préexiste à l'arrivée du LCME dans l'organisation, celle-ci peut également être considérée comme la plus pertinente pour l'indexation. Dans notre projet-pilote, la décomposition du domaine en sous-domaines nous a permis de nous affranchir de ce choix. En effet, le seul modèle des concepts commun à tous les sous-domaines est la notion de matière fissile. C'est donc cette notion qui a été choisie comme source pour l'indexation conceptuelle du fonds documentaire.

5.1.4.2.5. Du modèle du domaine à la création des règles de comportement du LMCE ?

Le modèle du domaine positionne et articule les principaux processus (physiques ou plus généralement comportementaux) qui régissent le fonctionnement d'un système. Ce modèle est la transcription diagrammatique directe des règles des systèmes experts utilisées en intelligence artificielle. Le nombre de paramètres déterminants pour l'application de ces règles est d'ailleurs la principale raison de la complexité de ce formalisme. Inversement, ce modèle permet l'extraction quasi-automatique de ces règles comportementales du système.

Notre objectif n'est plus ici de bâtir un système expert (nous avons souligné plus haut que le but du LMCE n'est pas le remplacement de l'opérateur mais l'accompagnement de celui-ci dans l'enchaînement des tâches qui composent son activité). Les règles que nous pouvons souhaiter extraire de la modélisation MKSM concernent le comportement du LMCE. Nous avons identifié au chapitre 4 trois types de règles de comportement de cet outil : les règles de formulation de la requête, de sélection des index et de comportement de l'interface.

La question qui se pose alors est de savoir s'il est possible d'établir une correspondance générique entre les règles comportementales d'un système physique modélisé et celles d'un outil d'accès à l'information spécifiquement conçu pour une organisation s'intéressant à ce système. En d'autres termes, comment et jusqu'à quel point les processus environnementaux de l'organisation doivent-ils influencer sur le comportement du LMCE ?

Nous n'avons pas d'élément de réponse à proposer pour cette question. Cela ne signifie pas nécessairement que le système modélisé et le LMCE sont totalement indépendants. Ils peuvent simplement être trop éloignés pour que de telles règles soient définies à un niveau générique. En particulier, le projet-pilote sur lequel nous avons travaillé s'intéresse à l'activité d'une organisation qui étudie un phénomène physique dans le cadre de l'activité d'autres organisations. Cette activité est donc suffisamment éloignée du phénomène physique

modélisé pour que celui-ci n'ait pas de répercussion sur le fonctionnement de cette organisation.

Toutefois, dans l'état actuel de réalisation du LMCE, nous sommes sceptique quant à l'opportunité de définir des règles de comportement de cet outil à partir du modèle du domaine MKSM. En effet, nous avons vu au chapitre 4 que les règles de formalisation de la requête sont uniquement fonction du langage de requêtes utilisé. De la même manière, les règles de sélection des index sont uniquement dictées par la structure arborescente du modèle de l'information et les règles de comportement de l'interface dépendent du profil de l'utilisateur. Nous voyons mal alors comment le fonctionnement synchronique du système modélisé peut influencer sur ces paramètres.

Dans la perspective d'une évolution de l'outil prenant en compte des paramètres plus fins du domaine modélisé, nous pouvons cependant indiquer quelques éléments quantitatifs permettant de relier ce domaine et un outil de présentation de ce domaine. D'une part, le nombre de processus et de relations entre ceux-ci sur le scénario des processus détermine le nombre de règles et donc la complexité du domaine. D'autre part, l'analyse du vocabulaire technique pour la reformulation des éléments des requêtes pourrait être automatisée. Dans cette hypothèse, le système de règles aurait également à gérer des règles de reformulation, plus proches du fonctionnement du domaine modélisé que les trois autres types de règles gérées actuellement.

5.1.4.2.6. *De l'utilité des autres modèles*

Nous avons identifié deux catégories d'autres formalismes : les formalismes spécifiques aux domaines modélisés et les autres formalismes génériques.

Concernant l'intérêt des formalismes spécifiques aux domaines modélisés, l'exemple des formalismes définis par Stéphane Picard dans le cadre d'un projet de gestion des connaissances appliqué aux codes de calcul scientifiques, nous semble représentatif ([Picard & al. 99]). Ces formalismes sont plus proches de ceux utilisés en génie logiciel, et donc de ceux dont sont familiers les destinataires de la modélisation. D'autre part, ils peuvent éventuellement être réutilisés dans des phases de conception logicielles ultérieures pour caractériser les éléments existants. Ces formalismes sont créés a priori pour être spécifiques à un domaine particulier. Leur application sort du cadre de l'accès à l'information et est également spécifique au domaine. Dans le cas du remplacement des formalismes MKSM « classiques » par des formalismes spécifiques au domaine, il convient de mener

l'argumentation ci-dessus en fonction du rôle de chacun de ces formalismes dans la modélisation. La même démarche doit être appliquée à la discussion suivante concernant le rôle de ces formalismes spécifiques pour l'accès à l'information avec la stratégie cumulative.

Les formalismes génériques, tels que le modèle des objets que nous avons utilisé dans le projet-pilote, les formalismes d'évolution de [Van Craeynest & al. 00] ou même les éléments textuels intégrés dans la modélisation sont destinés à tenir un « rôle » particulier dans la modélisation du domaine. En ce sens, ils ont une fonction documentaire spécifique au sein de cette modélisation et participent de la caractérisation du problème d'information des utilisateurs. De telles fonctions documentaires sont données en 5.1.2.4.3.

L'utilité de ces formalismes pour la conception du LMCE n'est a priori pas flagrante, puisque nous avons réussi à mener à bien le processus de conception de cet outil sans leur contribution. Toutefois, il est possible que cette utilité se manifeste concernant certains aspects de la conception de ce système que nous n'avons pas traités ici. Dans ce cas, la question de l'utilité de tels formalismes revient à s'interroger sur les aspects de l'accès à l'information non pris en compte dans ce travail. De tels aspects, comme notamment la prise en compte des outils de production de l'information, visant à repositionner l'accès à l'information dans la perspective plus globale du cycle de vie de cette information, sont évoqués dans les perspectives d'évolution de l'approche du bilan prospectif (paragraphe 5.2.3.).

5.1.4.3. Pertinence de l'approche de la gestion des connaissances dans le cadre plus général de sa réintroduction dans l'accomplissement de l'activité modélisée

La version du microscope ci-dessus fait état d'un axe de décomposition diachronique de l'activité modélisée. Les formalismes existants font état d'un fonctionnement présent et passé du système. Une question qui se pose est alors de savoir s'il est possible de prévoir une évolution future de ce système. La question du positionnement de la gestion des connaissances dans l'évolution diachronique du système modélisé se pose d'ailleurs depuis longtemps à la DIST du CEA, indépendamment de la problématique d'accès à l'information¹⁴⁶. Il s'agit dans cette partie d'étudier un aspect particulier de ce positionnement, qui est l'impact de la modélisation sur le système modélisé.

¹⁴⁶ Les modèles d'évolution de [Van Craeynest & al. 00] apportent un élément de réponse en ce qui concerne la modélisation de l'évolution diachronique passée. Une thèse sur la gestion des connaissances pour l'innovation

La modélisation de l'activité réalisée dans un projet de gestion des connaissances est la première étape vers la réalisation d'un LMCE. Nous avons vu au chapitre 4 que cette modélisation inclut l'identification des tâches d'accès à l'information avant l'introduction du LMCE dans l'activité. Cet outil a pour but de modifier la manière dont les individus accèdent à l'information dans l'accomplissement de l'activité modélisée. Par conséquent, les tâches d'accès à l'information – comme l'interrogation directe des bases – telles qu'elles sont décrites dans le LMCE ne sont plus valides dès l'entrée en fonction de cet outil.

Le même type de problème peut être rencontré quel que soit le type d'application de la modélisation des connaissances, dès lors que celle-ci est destinée à être réinjectée dans l'activité modélisée¹⁴⁷. La question que nous nous posons ici concerne la pertinence de l'approche du remplacement de solutions existantes par des outils incorporant une modélisation de l'activité dans laquelle ils ne figurent pas.

Nous proposons de traiter l'introduction de l'outil à concevoir dans le système modélisé comme une « perturbation » de ce système¹⁴⁸. Une « correction » possible est de modifier les diagrammes de connaissances en fonction de l'utilisation de l'outil. Le problème est alors l'utilité de la récursivité induite par cette modification. Par exemple, si l'utilisateur qui souhaite accéder à l'information avec le LMCE trouve dans un diagramme de tâches l'indication « Utiliser le LMCE », il est peu probable que cette indication lui soit utile pour son problème. Il s'agit alors de déterminer si les fonctionnalités de l'outil nécessitent de faire apparaître cette récursivité.

Si l'utilisation de l'outil ne fait appel qu'aux parties de la modélisation dans lesquelles il ne figure pas, nous proposons de ne pas faire figurer l'utilisation de l'outil dans la modélisation. Pour cela, il suffit de limiter la profondeur de la décomposition à un niveau de détails supérieur à la tâche mentionnant son utilisation.

est actuellement en cours pour explorer la prédictibilité de l'évolution d'un système grâce à la modélisation des connaissances.

¹⁴⁷ Ce problème n'est pas rencontré dans l'utilisation de la modélisation pour l'apprentissage, car cette modélisation n'est utilisée que durant la phase de formation, qui précède l'accomplissement de l'activité. La modélisation n'intervient donc pas directement dans cette activité. L'accès à l'information, l'utilisation des modèles dans l'activité de conception logicielle (comme les codes de calcul scientifiques) sont des composantes de l'activité du système.

¹⁴⁸ La théorie des perturbations utilisée en physique permet le calcul de l'impact d'une modification d'un système global, en postulant que la loi de fonctionnement de ce système reste valable au voisinage du point de fonctionnement considéré. Elle procède alors par l'application d'un coefficient correcteur dans le développement limité de la fonction modélisant le phénomène étudié.

Si la tâche supplantée par l'outil à concevoir nécessite une description dans la modélisation, l'utilisation de l'outil doit apparaître dans cette modélisation. Dans ce cas, il est logique de faire figurer les tâches prescrites par les concepteurs du système ([Tricot & Nanard 98]). La modélisation est alors utilisée à la fois pour l'accomplissement du but de l'utilisateur et comme un guide d'utilisation de l'outil. La difficulté est alors de déterminer ces tâches en collaboration avec les experts, puisque le chargé de projet n'est pas supposé compétent dans l'accomplissement de ces tâches. La problématique de conception de l'outil est alors compromise par le fait que la modélisation est nécessaire à la définition de ses fonctionnalités, mais que ces fonctionnalités ne peuvent être définies en l'absence de la modélisation. Il convient de résoudre les problèmes posés par cette récursivité. Une méthode employée peut être celle des approximations successives. Il s'agit de définir des fonctionnalités en fonction de la modélisation, de les incorporer à la modélisation, et de boucler jusqu'à ce que le système converge. Cette méthode de conception, du fait qu'elle nécessite l'intervention des experts, s'apparente à une démarche de conception participative (cf. [Mackay & al. 98]). Toutefois, il convient de s'interroger sur le coût d'une telle méthode de conception pour l'organisation, comparément à celui d'une approche « classique ».

Par conséquent, le problème de la pertinence de la réintroduction de la gestion des connaissances dans l'activité de l'organisation consiste à déterminer si la tâche supplantée par l'outil nécessite d'apparaître dans la modélisation. Le cas échéant, la pertinence d'une approche réinjectant la modélisation dans l'activité modélisée n'est pas avérée. Le travers de cette approche est que l'introduction du système rend obsolète la modélisation réalisée. La correction de ce travers peut nécessiter des efforts importants en regard d'autres solutions. Dans le cas contraire, la présence d'éléments de modélisation de l'activité de l'outil n'a pas d'incidence sur sa conception¹⁴⁹. Nous considérons que le LMCE se trouve dans ce cas, car le détail de la tâche d'accès à l'information n'apporte pas de précision quant à la nature du problème d'information de l'utilisateur.

En ce sens, nous pensons qu'une approche par la gestion des connaissances ne permet pas de concevoir des outils destinés à tous les domaines d'application des activités modélisées. L'intérêt de cette approche est cependant réel pour des applications connexes à l'activité des organisations. C'est le cas de l'accès à l'information pour les organisations dont ce n'est pas

¹⁴⁹ Cela ne signifie pas qu'elle ne doit pas être modélisée. Par exemple, nous avons signalé la nécessité de modéliser les tâches d'accès à l'information pour la conception du LMCE, qui sont essentielles à l'identification des critères d'indexation conceptuelle.

l'activité principale. La conception d'un LMCE est par exemple plus pertinente pour le SEC que pour une bibliothèque.

5.2. Bilan prospectif

Dans ce chapitre, nous avons discuté plusieurs aspects de notre démarche d'accès à l'information scientifique et technique fondée sur la modélisation de l'activité. Cette discussion est menée en l'absence des tests de validation ergonomique du prototype.

Dans une première partie, nous indiquons les apports et les limites de notre démarche d'accès à l'information par recontextualisation et de la stratégie d'accès à l'information cumulative associée. Il apparaît que la pluridisciplinarité de cette approche offre une certaine richesse en termes de modélisation et de possibilités. En revanche, elle est coûteuse à mettre en place.

La deuxième partie discute, d'un point de vue ergonomique, les apports de la modélisation des connaissances pour l'accès à l'information. Il ressort que la gestion des connaissances se positionne comme une carte contextuelle complexe mais potentiellement efficace. L'utilisation de cette stratégie nécessite une formation à la méthode de présentation de la modélisation utilisée. Moyennant une telle formation, la richesse des relations proposées par cette méthode permet de diminuer le coût cognitif de la tâche d'accès à l'information. Cela est notamment réalisé en proposant des « raccourcis cognitifs » qui permettent l'identification par le système du contexte de la tâche de recherche d'information de l'utilisateur. La stratégie d'accès à l'information cumulative est discutée d'un point de vue essentiellement ergonomique. L'objectif est d'évaluer la pertinence de chacun des formalismes pour la catégorisation du problème d'information d'un utilisateur.

Dans la troisième partie, nous rappelons l'apport de l'ingénierie de l'information pour notre approche. Outre l'utilisation d'un moteur de recherche en texte intégral pour l'accès à l'information, qui apparaît comme évident pour ce type d'applications, il s'agit de discuter le paradigme choisi de l'indexation conceptuelle des documents. Celui-ci s'avère nécessaire pour la mise en relation des modèles du contexte d'utilisation de l'information et des sources documentaires.

Enfin, dans la quatrième partie, nous intéressons à la discipline qui est à la base de notre approche : la gestion des connaissances. Un examen plus approfondi des multiples implications de cette discipline dans ce travail est nécessaire. L'utilisation de la modélisation des connaissances est évaluée sous trois aspects.

- Tout d'abord, les fondements de la méthode MKSM par rapport à notre problématique sont discutés. Nous mettons notamment en lumière la faible pertinence du paradigme sémiotique sur lequel s'appuie cette méthode dans le cas d'une modélisation d'un domaine à plusieurs niveaux. Le repositionnement des modèles dans le processus plus général de l'accès à l'information permet d'aboutir à un référentiel hypermédia pour le partage de l'information orienté utilisation.
- Puis, nous montrons que les différents formalismes permettent d'utiliser la méthode MKSM comme une méthode de génie logiciel, en proposant des représentations adéquates pour chacune des phases de la conception. En tant que référentiel de conception de l'outil d'accès à l'information, la méthode MKSM propose un cadre permettant d'unifier les étapes de la modélisation dans les différentes disciplines. Les coûts impliqués par l'acquisition des données nécessaires à la mise au point des modèles sont ainsi réduits.
- Enfin, la finalité de la méthode est discutée. La question posée est relative à la perturbation impliquée par l'incorporation de la modélisation d'un système dans sa propre activité. Cette question ne se pose pas pour le LMCE, car cet outil ne permet pas un parcours « récursif » de la modélisation. Pour d'autres applications, cette question reste toutefois ouverte.

Le paragraphe ci-dessus clôt la discussion sur le travail réalisé pendant cette thèse. Les suivants indiquent le travail qui reste à accomplir et les pistes que nous considérons comme utiles à explorer à la suite de ce travail.

5.2.1. *Processus d'évaluation et de validation*

La méthode MKSM suppose des interviews d'experts des unités où seront installés les LMCE. Nous sommes dès lors directement au contact des futurs utilisateurs de ces systèmes. La démarche de rédaction, validation et amélioration des diagrammes, appliquée à l'identification des besoins en information des experts et des fonctionnalités du futur système¹⁵⁰, nous permet donc de supposer valides le contenu et la structuration de la modélisation des connaissances obtenue sur le domaine. D'après la méthode proposée au chapitre 4, la phase de modélisation permet également d'obtenir les modèles des utilisateurs et du SI, qui sont également soumis aux experts du domaine pour validation. Dans le projet-pilote sur la sûreté-criticité, les différentes collections de diagrammes de connaissances correspondant aux sous-domaines présentés dans le prototype ont été validés par le SEC.

Nous avons de plus reçu un accord de principe concernant les particularités des différents profils utilisateurs (niveaux d'expertise, modes de filtrage) proposés ainsi que sur le modèle du système d'information réalisé.

Au début du processus d'évaluation et de validation, nous pouvons donc présupposer une certaine validité des modèles conceptuels sur lesquels s'appuie le prototype. Il s'agit dans cette partie de définir ce processus qui doit débiter dès la fin de la phase de conception, et qui doit avoir lieu en coopération avec l'équipe du projet MerLIn (Méthodes pour l'Ergonomie des Logiciels Interactifs) de l'INRIA. Comme nous l'avons vu au chapitre 2, cette phase d'évaluation comporte deux objectifs : une validation de l'utilisabilité du prototype et une évaluation de son utilité

5.2.1.1. Inspection ergonomique

Le processus de validation de l'utilisabilité du prototype est une démarche itérative qui s'appuie sur les résultats d'une évaluation pour déboucher soit sur une modification du prototype, soit sur sa validation (cf. paragraphe 3.2.2.2.2.6). Cette évaluation de l'utilisabilité du prototype a pour objet de déceler d'éventuels manquements aux règles ergonomiques de conception. Il s'agit d'une part de vérifier si ces règles, passées en revue pendant le processus de conception, ont été correctement prises en compte, et d'autre part, d'évaluer l'adéquation du prototype réalisé avec l'outil décrit au chapitre 4. L'évaluation de l'utilisabilité fait appel à une étape d'inspection ergonomique et une étape de tests utilisateurs (voir paragraphe 2.3.4.4).

L'inspection ergonomique est réalisée par des évaluateurs experts en ergonomie. Elle s'appuie sur leur connaissance des règles de conception et des critères d'évaluation ergonomiques pour identifier les éventuelles erreurs de conception qui perturbent la compréhension globale du fonctionnement de l'interface du LMCE. L'intérêt de cette méthode est qu'elle ne nécessite pas la participation des utilisateurs.

Concrètement, la stratégie d'évaluation du prototype est fondée sur des scénarios d'utilisation. Chacun de ces scénarios permet de tester les tâches relatives aux fonctionnalités décrites au paragraphe 4.2.5.1. Ces scénarios sont décrits dans le tableau ci-dessous.

Profils des	Fonctionnalités	Scénarios
-------------	-----------------	-----------

¹⁵⁰ Les différents avantages et inconvénients de la modélisation conjointe des connaissances des experts du domaine et des utilisateurs du LCE sont détaillés dans [Médini 99].

utilisateurs	à tester	
Tous types de profils	Création d'un profil	<ul style="list-style-type: none"> • Sélection de la fonctionnalité « Création d'un profil » sur l'écran d'introduction • Renseignement (correct et erroné) de toutes les informations demandées en fonction des caractéristiques des utilisateurs • Envoi du formulaire • Fin : récupération du feedback
Tous types de profils	Connexion	<ul style="list-style-type: none"> • Renseignement (correct et erroné) des champs « nom » et « prénom » sur l'écran d'introduction • Envoi du formulaire • Fin : récupération du feedback
Tous types de profils	Modification d'un profil	<ul style="list-style-type: none"> • Connexion • Sélection du bouton « Options » dans le cadre de commandes utilisateur • Sélection du bouton « Modification de votre profil » • Modification (correcte et erronée) de certaines caractéristiques du profil courant • Envoi du formulaire • Fin : récupération du feedback
Novice	Navigation simple	<ul style="list-style-type: none"> • Connexion en tant qu'utilisateur novice • Lecture du diagramme présenté et de la carte conceptuelle associée • Sélection et activation d'un lien à partir d'un élément du diagramme • Sélection et activation d'un lien à partir de la carte conceptuelle présentée • Boucle : répétition des trois dernières étapes
Tous types d'apprenants	Sélection d'éléments des diagrammes pour l'accès à l'information	<ul style="list-style-type: none"> • Connexion en tant qu'utilisateur apprenant • Navigation dans la collection de diagrammes • Clic sur un élément du diagramme présenté • Sélection du lien « Ajouter au formulaire de recherche » • Boucle : répétition des trois dernières étapes
Tous types d'apprenants	Composition et envoi d'une requête avec la stratégie cumulative	<ul style="list-style-type: none"> • Sélection d'éléments des diagrammes pour l'accès à l'information • Modification de la position des « radio buttons » du caddie à concepts • Modification des valeurs des champs texte du caddie à concepts • Affichage / modification de la requête • Masquage de la requête • Envoi de la requête • Fin : récupération du feedback
Expert ingénieur	Interrogation directe	<ul style="list-style-type: none"> • Composition d'une requête contenant un seul terme (« INB »), plusieurs termes compatibles (« INB ET uranium ») et plusieurs termes incompatibles (« INB ET NON INB ») • Envoi de la requête

		<ul style="list-style-type: none"> • Récupération du feedback • Boucle : répétition des trois dernières étapes
Expert chercheur	Interrogation par champs	<ul style="list-style-type: none"> • Composition d'une requête par renseignement de plusieurs champs : auteur, type de document, date, origine et texte • Envoi de la requête • Récupération du feedback • Boucle : répétition des trois dernières étapes avec des valeurs des champs différentes

Tableau 8. Scénarios d'utilisation de la phase d'inspection ergonomique.

Si les résultats de l'inspection ergonomique du système font état de problèmes de conception, une nouvelle itération de conception et d'évaluation est nécessaire pour les corriger. Selon le type de problème détecté, cette « re-conception » du prototype peut nécessiter un « simple » retour à la phase de programmation ou une redéfinition des tâches « élémentaires »¹⁵¹ des utilisateurs, également suivie d'une reprogrammation de la (des) tâche(s) modifiée(s).

Cependant, une inspection ergonomique n'est a priori pas suffisante pour évaluer tous les aspects de l'utilisabilité du prototype. Deux points sont encore à évaluer avant de pouvoir statuer sur cette utilisabilité.

D'une part, la pertinence de l'emploi de certains termes spécifiques ne peut être déterminée par un évaluateur extérieur au domaine. Dans ces cas, le recours à des tests impliquant les utilisateurs finaux est nécessaire. Cependant, nous avons vu que les éléments spécifiques au domaine (i.e. les trois modèles conceptuels) sont supposés valides. De plus, la portée de cette évaluation de l'utilisabilité du prototype se restreint aux éléments génériques de cet outil. L'évaluation du contenu serait a priori plus du ressort de l'évaluation de l'utilité que de l'utilisabilité du LMCE.

D'autre part, des données provenant de l'observation des utilisateurs peuvent être requises pour évaluer certains choix de conception. Par exemple, pour ordonner de façon optimale les différents liens dans le multiplexeur, la donnée des fréquences respectives d'activation de ces liens peut être importante. Ce type de paramètres ne peut être évalué hors de la phase de tests utilisateurs présentée ci-dessous.

¹⁵¹ L'évaluation de l'utilisabilité du processus ne porte que sur les tâches déterminant les types d'interaction avec le système. Un exemple de ce type de tâche est la tâche d'interaction avec un multiplexeur de liens. Celle-ci peut être remplacée par une tâche de sélection des différents liens dans un « pop-up menu » accessible par un clic droit de la souris. La pertinence des tâches « fondamentales » du système, qui sont liées à l'approche, sont évaluées par l'utilité du prototype.

5.2.1.2. Tests utilisateurs

La phase de tests utilisateurs peut servir à prolonger l'évaluation de l'utilisabilité du prototype en fonction des résultats de la phase précédente. Dans ce cas, la détermination des fonctionnalités à tester et la modalité des tests incombe aux experts ayant mené la phase d'inspection ergonomique. Comme précédemment, si ces tests mettent au jour la nécessité d'une nouvelle itération de conception et d'évaluation, une version 2 du prototype devra être programmée. L'objet de ce paragraphe est de présenter les tests utilisateurs qui doivent être menés pour évaluer l'utilité du système.

L'évaluation de cette utilité a pour but de mesurer la pertinence de notre démarche d'accès à l'information par recontextualisation. Le LMCE propose en plus de la stratégie cumulative, plusieurs stratégies d'accès à l'information analytiques, dont la pertinence n'a pas à être validée. Ces stratégies peuvent alors nous servir de référents pour l'évaluation de la stratégie cumulative. Il ne s'agit pas ici d'évaluer les performances absolues du LMCE en tant qu'outil d'accès à l'information, mais de comparer la pertinence de la stratégie cumulative à celle des deux stratégies analytiques proposées.

Pour cela, la phase de tests utilisateurs doit permettre d'obtenir des données quantitatives sur les performances en termes de rappel et de précision (cf. paragraphe 2.2.2.3) de la recherche d'information et sur le temps mis par les utilisateurs pour réaliser les différentes tâches d'accès à l'information avec chaque stratégie¹⁵². Les données obtenues par chaque stratégie sont alors moyennées sur toutes les recherches effectuées¹⁵³, puis normalisées en fonction de celles obtenues par la stratégie de référence (i.e. interrogation directe pour les sous-domaines d'ingénierie et interrogation par champs pour les sous-domaines de recherche).

Par exemple, si l'on nomme r_c , p_c et t_c les moyennes de rappel, de précision et de temps donnés par la stratégie cumulative pour chaque sous-domaine, r_{id} , p_{id} et t_{id} ces mêmes données pour la stratégie d'interrogation directe et r_{ic} , p_{ic} et t_{ic} pour l'interrogation par champs, la mesure comparative de ces trois valeurs pour le sous-domaine des études de conception pour les INB est :

$$r_{INB} = r_c / r_{id} ; p_{INB} = p_c / p_{id} ; t_{INB} = t_c / t_{id}.$$

¹⁵² Le temps de réponse du système n'est en revanche pas mesuré, car il ne dépend a priori pas de la stratégie utilisée pour générer la requête (le contrôle ActiveX déclenchant le plug-in d'interrogation Acrobat Search® est le même pour les trois stratégies).

¹⁵³ L'échec d'une démarche d'accès à l'information (abandon du sujet avant l'envoi de la requête) doit être comptabilisé comme une requête n'ayant donné aucun résultat pertinent (rappel = précision = 0).

De même, pour le sous-domaine des expériences critiques, on obtient :

$$r_{\text{exp}} = r_c / r_{ic} ; p_{\text{exp}} = p_c / p_{ic} ; t_{\text{exp}} = t_c / t_{ic}.$$

Le fait de réaliser une évaluation comparative des stratégies d'accès à l'information ne nécessite pas de connaître les performances du système dans des conditions réelles d'utilisation (i.e. pour une interrogation de la totalité du système d'information de l'unité). La seule contrainte est d'utiliser la même base de tests pour l'évaluation de la stratégie analytique de référence et de la stratégie cumulative. En conséquence, les tests utilisateurs ne doivent pas nécessairement avoir lieu dans l'environnement de travail des utilisateurs et peuvent être menés en laboratoire. Les modalités de cette phase de tests sont donc à discuter entre l'unité destinataire du système (le SEC) et les évaluateurs (le projet MErLIn). D'un point de vue logistique, les tests en laboratoire sont moins complexes à mettre en place. En revanche, les tests sur site monopolisent les sujets moins longtemps, et permettent donc, à coût égal pour le SEC, une expérimentation sur un plus grand nombre de sujets.

Une autre donnée dont nous ne sommes pas maître est le nombre d'utilisateurs qui participeront à l'évaluation. Les utilisateurs qui nous intéressent sont les destinataires de notre stratégie cumulative (i.e. des apprenants et des experts non familiarisés avec les outils de recherche d'information). Le nombre de participants dépend d'une part de l'unité, qui peut rechigner à se séparer d'éléments importants pour son fonctionnement¹⁵⁴, et d'autre part du nombre d'individus répondant à ces caractéristiques dans l'unité et acceptant de se soumettre à ces tests.

Quoi qu'il en soit, ces tests consistent à demander aux sujets d'effectuer des séries de deux démarches de recherche d'information, l'une avec la stratégie cumulative, et l'autre avec la stratégie analytique de référence. Des exemples de scénarios nécessitant la réalisation de ces tâches sont proposés dans le tableau ci-dessous. Ils ne sont qu'indicatifs, et les sujets peuvent choisir de réaliser une tâche d'accès à l'information différente – et sans doute plus réaliste – s'ils le désirent, du moment que les objectifs de la recherche sont les mêmes pour les deux stratégies.

Domaine	Stratégies à comparer	Documents recherchés
INB	Cumulative	Recherche d'un type de document (rapport d'étude de criticité),

¹⁵⁴ Dans le cas où les tests se passeraient en laboratoire et où le déplacement des sujets les rendrait indisponibles pour une longue période.

	vs directe	d'une configuration de matière fissile (Uranium 235 faiblement enrichi en solution) et d'une partie d'INB (dissolveur).
Expériences	Cumulative vs par champs	Recherche d'un type de document (rapport d'expérience critique) et d'une configuration de m.f. (plutonium sous forme de crayons émettant des neutrons de spectre intermédiaire).

Tableau 9. Scénarios de test des différentes fonctionnalités du LMCE.

5.2.1.3. Questionnaire de satisfaction

En complément de ces données, il est prévu de proposer aux utilisateurs un questionnaire pour leur permettre d'indiquer leurs commentaires subjectifs sur le prototype. Ce questionnaire doit permettre d'obtenir un retour d'expérience à plus long terme sur la satisfaction des utilisateurs et sur les stratégies qu'ils ont choisi d'adopter. Il permet également d'obtenir une évaluation subjective mais absolue de l'utilité du système. L'idée est de permettre une utilisation du système pendant un temps suffisamment long¹⁵⁵ et de recueillir ensuite les impressions globales sur l'accès à l'information par recontextualisation. Pour cela, il s'agit de proposer aux utilisateurs de remplir un questionnaire de satisfaction.

L'objectif de cette phase est d'arriver à analyser finement l'opinion d'un grand nombre d'utilisateurs sur les différents aspects du prototype. Un questionnaire peut permettre de poser des questions suffisamment précises pour différencier l'évaluation des différents aspects du prototype. En particulier, les aspects spécifiques au domaine peuvent ainsi être dissociés des aspects méthodologiques de l'approche.

En pratique, ce questionnaire comporte deux parties. D'une part, une série de questions fermées doit permettre de déterminer l'opinion que les utilisateurs se seront faite du prototype. Il s'agit de déterminer leur fréquence d'utilisation du LMCE, les stratégies et les fonctionnalités qu'ils utilisent et leur satisfaction générale. Cette partie pourra être réalisée sous la forme d'une série d'affirmations, auxquelles les utilisateurs devront répondre par vrai ou par faux (cf. questionnaire WAMMI, [Bastien & al. 98] pp. 152-154). D'autre part, il contiendra des questions ouvertes, permettant notamment aux utilisateurs de proposer des suggestions de modifications à apporter au prototype.

¹⁵⁵ Un mois nous paraît a priori un délai raisonnable.

5.2.1.4. Analyse des données recueillies et bilan du processus d'évaluation

L'analyse des données ainsi recueillies doit permettre d'évaluer plusieurs aspects de ce travail. L'objectif est de tenir compte à la fois des données quantitatives issues des tests et des données subjectives fournies par le questionnaire pour évaluer les points suivants :

- Utilisabilité du prototype : malgré les différentes étapes d'évaluation du prototype ci-dessus, nous ne pouvons pas exclure des problèmes résiduels liés à l'utilisabilité du système. De tels problèmes pourraient par exemple être mis à jour par le retour d'expérience des utilisateurs et signalés dans les réponses aux questions ouvertes du questionnaire.
- Validité des modèles : si les modèles conceptuels du domaine sont supposés valides au début de l'évaluation, ils peuvent cependant être remis en cause par plusieurs facteurs. Un faible indice de satisfaction des utilisateurs, toutes stratégies confondues, peut nous conduire à nous interroger sur la validité des index conceptuels. La navigation fréquente d'utilisateurs correspondant à un profil donné dans un autre sous-domaine que celui qui les intéresse peut invalider le découpage des sous-domaines dans la modélisation ou dans les profils utilisateurs.
- Validité de l'approche : à ce stade, il serait périlleux de tenter d'en définir les modalités, qui seront dépendantes des résultats obtenus. En effet, dans le cas idéal où tous les résultats quantitatifs obtenus (temps, rappel, précision) pour les différentes stratégies sont conformes à nos attentes (pour chaque type d'utilisateurs, les meilleurs résultats sont obtenus pour la stratégie qui lui correspond), nous pourrions statuer sur la validité de notre approche d'accès à l'information par recontextualisation pour les types d'utilisateurs correspondants. Dans tous les autres cas, il faudra procéder à une analyse plus fine des données recueillies pour les situations où ces données ne sont pas conformes à nos attentes et tenter de trouver les raisons des problèmes soulevés. En tout état de cause, ces conclusions devront être pondérées par les données subjectives issues du questionnaire, reflétant la satisfaction générale des utilisateurs vis-à-vis de chacune des stratégies d'accès à l'information proposées.

À l'issue de ce processus d'évaluation, nous espérons tirer des enseignements sur les points forts et les points à améliorer du prototype et de notre approche. Nous espérons que ces enseignements pourront être utilisés soit pour corriger certains des points à améliorer, soit pour réaliser des améliorations que les utilisateurs nous auront suggérées.

5.2.2. Perspectives d'évolution de l'outil

Plusieurs directions n'ont pas été explorées dans cette thèse, mais nous semblent toutefois intéressantes pour prolonger ce travail. Les pistes indiquées ici concernent des évolutions envisageables du LMCE. À court terme, il s'agit de faire évoluer les solutions techniques pour l'utilisation d'un moteur de règles permettant de regrouper les aspects dynamiques du fonctionnement du LMCE. À plus long terme, il s'agit d'élargir le spectre des types d'informations accessibles ainsi que le langage d'interrogation utilisé par le LMCE.

5.2.2.1. Perspectives offertes par l'utilisation d'un moteur de règles

Nous avons vu au chapitre 4 que le fonctionnement dynamique du LMCE est géré par des scripts incorporés aux différents éléments de l'interface. À terme, il est possible d'envisager la suppression des scripts de ces composants de l'interface. Ces scripts ne font en fait qu'appliquer des règles dans le contexte des éléments de l'interface qu'ils gèrent. Le comportement de l'interface serait alors directement commandé par le moteur de règles. Le LMCE gagnerait sans doute également en cohérence et en facilité de maintenance.

L'utilisation d'un moteur de règles (par exemple en Prolog) pourrait servir à permettre à l'utilisateur de spécifier ses préférences de visualisation sous forme de règles (exactement comme dans les feuilles de style XSL). Dans la même optique, un tel outil pourrait encore améliorer la prise en compte des paramètres spécifiques de l'utilisateur en lui permettant de spécifier des règles de filtrage personnalisées. Il pourrait par exemple spécifier qu'il s'intéresse à certains sous-domaines mais pas à d'autres ou définir des modes de filtrage spécifiques en fonction des termes de la requête.

En outre, la gestion des relations sémantiques entre les éléments de la modélisation MKSM serait facilitée. Nous pouvons par exemple imaginer la prise en charge, de la pose à la navigation hypermédia, des liens entre les diagrammes de connaissances. Une telle fonctionnalité permettrait alors le calcul dynamique de ces liens et la suppression du diagramme en tant qu'entité informationnelle. Les éléments des diagrammes actuels pourraient alors être repositionnés dynamiquement en fonction de règles ergonomiques spécifiques à la visualisation de ces éléments sur écran, ce qui bénéficierait grandement à la lisibilité de la modélisation.

Toutefois, l'utilisation de règles pour la gestion de toutes les fonctionnalités du système accroîtrait considérablement la complexité de ce système. Il serait alors raisonnable de

s'interroger sur l'impact de cette amélioration en termes de performance de la recherche et de rapidité d'exécution.

5.2.2.2. Vers une indexation conceptuelle multi-supports par le LMCE

Jusqu'ici, nous ne traitons que de l'information textuelle sur support numérique. Le LMCE n'est pas conçu pour permettre l'accès à des éléments textuels disponibles sur support papier ou à l'information multimédia. Une manière de permettre l'accès aux documents papiers est de disposer d'une base de références au format numérique, qui est alors traitée – presque – comme les autres bases documentaires¹⁵⁶.

L'essor des nouvelles technologies de l'information et de la communication laisse entrevoir l'apparition de plus en plus fréquente dans les systèmes d'information des organisations d'éléments de types non textuels. Les images, les documents sonores et la vidéo sont appelés, sinon à remplacer, du moins à coexister avec l'information textuelle dans les systèmes d'information des organisations modélisées. C'est pourquoi il nous paraît important d'envisager l'indexation de ces types de sources d'informations dans le LMCE. Pour cela, nous indiquons quelques travaux menés en ce sens dans des domaines attenants.

Des travaux dans le domaine de l'indexation de documents vidéo pour la gestion des connaissances ont été entrepris à la DIST du CEA par Philippe Bigeon. Ils visent à employer ce support pour le recueil de l'expertise durant la phase de modélisation. Concernant l'indexation sonore, l'exemple des travaux conjoints des équipes travaillant sur la reconnaissance automatique de la parole et sur le traitement automatique du langage naturel du Laboratoire d'Informatique d'Avignon nous semble être une piste prometteuse. Concernant l'indexation, nous pouvons notamment citer les travaux sur les bases audiovisuelles menés à l'INA par l'équipe de Bruno Bachimont (cf. [Dechilly & al. 99]).

5.2.3. Perspectives d'évolution de l'approche d'accès à l'information par recontextualisation

Une piste intéressante à explorer pour le prolongement de l'approche de l'accès à l'information par recontextualisation est le passage de l'interrogation booléenne à l'interrogation en langage naturel. L'objectif est de permettre au LMCE d'aider l'utilisateur à formuler des requêtes directement en langage naturel. Si des travaux ont été réalisés dans la

¹⁵⁶ Voir par exemple les travaux de Gary Perlman sur la conception de bases bibliographiques.

littérature pour la formulation graphique de requêtes documentaires booléennes (cf. Golovchinsky & Chignell 93]), nous n'en avons pas trouvé s'appliquant à la formulation de requêtes en langages naturel.

Il nous semble cependant intéressant d'explorer les relations entre les processus de formulation de ces requêtes en langage naturel, plus riches que les langages booléens, et la navigation dans les modèles de connaissances. L'idée est d'arriver à une spécification plus précise du besoin d'information par l'utilisation de structures grammaticales complexes. Il s'agit notamment d'évaluer l'intérêt de permettre à l'utilisateur une « articulation grammaticale » des éléments sélectionnés dans les diagrammes. Cette évaluation serait à réaliser en termes d'utilisabilité de l'interface d'accès à l'information et de pertinence des réponses données par le moteur de recherche. L'utilisation de prépositions pour spécifier des compléments circonstanciels nous paraît une voie intéressante.

Techniquement, cette voie requiert en premier lieu l'utilisation d'un moteur de recherche en langage naturel, comme SPIRIT. De plus, le formulaire d'interrogation doit être modifié. Une solution est de remplacer les « radio buttons » permettant de choisir un opérateur booléen par des listes déroulantes (« combo boxes ») permettant la sélection des prépositions appropriées. Dans le cadre du projet-pilote, la requête booléenne « définir une expérience en criticité ET optimiser ET dissolvant ET plutonium » pourrait alors être remplacée par « définir une expérience en criticité ET optimiser POUR dissolvant AVEC plutonium ». Nous sommes ici encore loin d'une réelle requête en langage naturel, mais moyennant le système de règles « amélioré » décrit plus haut, nous pensons qu'une génération approximative des articles peut être réalisée à peu de frais. De plus, nous ne savons pas si la présence des mots vides est réellement nécessaire à l'analyse de la syntaxe de la requête avant leur suppression dans les moteurs de recherche en langage naturel comme SPIRIT.

Bibliographie

- [Bastien & al. 98] : Bastien, J. M. C., Leulier C., Scapin D-L., *"L'Ergonomie des sites web"*, in *Créer et Maintenir un service web*, J-C. Le Moal & B. Hidoine Eds., ADBS, Paris, pp. 111-173, 1998.
- [Bernstein 90] : Bernstein M., *"Hypertext and technical writing"*, in proceedings of the ECHT'90 Course, Versailles, Novembre 1990.
- [Bernstein 93] : Bernstein M., *"Enactment in information farming"*, in proceedings of HyperText'93 Conference, ACM Press, pp. 242-249, 1993.
- [Chaillot & Ermine 97] : Chaillot M., Ermine J-L., *"Le livre de connaissances électronique"*, Document numérique, 1(1), 75-98, 1997.
- [Crampes & Ranwez 00] : Crampes M., Ranwez S., *"Ontology Supported and Ontology-Driven Conceptual Navigation on the World Wide Web"*, in proceedings of the ACM HyperText'2000 Conference, pp. 191-199, 2000.
- [Dechilly & al. 99] : Dechilly T., Auffret G., Brunie V., Bachimont B., *"Représentation et structuration des connaissances pour les bibliothèques audiovisuelles"*, Document Numérique, vol. 3, n° 3-4, Hermès, pp. 195-214, Paris, 1999.
- [de Vriès & de Jong 97] : de Vriès E., de Jong T., *"Using information systems while performing complex tasks: an example from architectural design"*, International Journal of Man-Machine Studies, vol. 46, n° 1, pp. 31-54, 1997.
- [Ermine & al. 96] : Ermine J-L., M. Chaillot, P. Bigeon, B. Charreton, D. Malavieille: *"MKSM, a method for knowledge systems management"*. in Knowledge Management, Organization, Competence and Methodology, Advances in Knowledge Management, Volume 1, Jos F. Schreinmakers Ed., pp. 288-302, Ergon Verlag, 1996.
- [Ganascia 93] : Ganascia J-G., *"Que Sais-Je ? Les Sciences Cognitives"*, Dunod, Paris, 1993.
- [Golovchinsky & Chignell 93] : Golovchinsky G., Chignell M., *"Queries-R-Links: Graphical Markup for Text Navigation"*, in proceedings of the ACM INTERCHI'93, Conference on Human Factors in Computing Systems, pp. 454-460, 1993.
- [Iwayama & Tokunaga 95] : Iwayama M., Tokunaga T., *"Cluster-Based Text Categorization: A Comparison of Category Search Strategies Text Categorization"*, in proceedings of the 8th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval p.273-280, 1995.
- [Lainé-Cruzel & al. 96] : Lainé-Cruzel S., Lafouge T., Lardy J-P. & Ben Abdallah N., *"Improving information retrieval by combining user profile and document segmentation"*, Information Processing and Management, Elsevier Science Ltd, Vol. 32, No. 3, pp. 305-315, 1996.
- [Mackay & al. 98] : Mackay W.E., Fayard A-L., Frobert L., Médini L., *"Reinventing the familiar: Exploring an augmented reality design space for air traffic control"*, in Proceedings of the ACM CHI'98, Conference on Human Factors and Computing Systems, ACM Press, pp. 558 - 565, Los Angeles, CA, 1998.
- [Médini 97] : Médini L., *"Conception et Réalisation d'un Livre de Connaissances Électronique sous forme de Document Hypermédia"*, Rapport de projet de DEA, Orsay et INSTN, Sept. 1997.
- [Nanard & Nanard 91] : Nanard J., Nanard M., *"Using Structured Types to incorporate*

- Knowledge in Hypertexts*", in proceedings of the Hypertext'91 Conference, ACM Press, pp. 329-334, 1991.
- [Picard & al. 99] : Picard S., Ermine J-L., Scheurer B., *"Gestion des connaissances pour des grands logiciels de Calcul scientifique"*, Plate-forme AFIA, IC'99 : Ingénierie des Connaissances, pp. 171-180, juin 1999, disponible en ligne : <http://www.afia.polytechnique.fr/plate-forme/Actes/ic18.pdf>.
- [Prié 01] : Prié Y., *"Sur la piste de l'indexation conceptuelle de documents. Une approche par l'annotation"*, dans le numéro spécial de la revue Document Numérique sur l'indexation, sous la direction de J-M Jolion, Document Numérique, Vol 4, n°1-2/2000, Hermès Science Publication, pp. 11-35, 2001.
- [Ranwez 00] : notes personnelles prises lors de la soutenance de thèse de Sylvie Ranwez, Décembre 2000.
- [Rouet & Tricot 98] : Rouet J-F., Tricot A., *"Chercher de l'information dans un hypertexte : vers un modèle des processus cognitifs"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 57-74, 1998.
- [Roussey & al. 99] : Roussey C., Calabretto S., Pinon J-M., *"État de l'art en indexation et recherche d'information"*, Document Numérique, Hermès, vol. 3, n° 3-4, pp. 121-149, Paris, 1999.
- [Salton & Mc Gill 83] : Salton G., Mc Gill M-J. *"Introduction to modern information retrieval"*, McGraw-Hill, pp. 157-198, 1983.
- [Tricot & Nanard 98] : Tricot A., Nanard J., *"Un point sur la modélisation des tâches de recherche d'information dans le domaine des hypermédias"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 35-56, 1998.
- [Van Craeynest & al. 00] : Van Craeynest J-M., Charlot J-M., Malavieille D., Bigeon P., Chaillot M., *"Traçabilité des connaissances avec MKSM"*, Actes du congrès scientifique NîmesTIC'2000, Nîmes, pp. 12-22, Septembre 2000.
- [Vogel 88] : Vogel. C., *"Génie cognitif"*, Collection Sciences Cognitives, Masson, Paris, 1988.
- [Zizi & Beaudoin-Lafon] : Zizi M., Beaudoin-Lafon M., *"Accessing Hyperdocuments through Interactive Dynamic Maps"*, in proceedings of the ACM ECHT'94, European Conference on Hypermedia Technologies, pp. 126-135, septembre 1994.

Chapitre 6. Conclusion

Cette thèse s'intéresse à la problématique de l'accès à l'information dans le contexte de l'exercice d'une activité à dominante scientifique et technique. Notre approche pluridisciplinaire concerne la définition d'une stratégie spécifique d'accès à l'information pour les individus concernés par cette problématique. Un outil informatique est proposé pour mettre en œuvre cette approche et tester sa validité sur le plan de l'efficacité de l'accès à l'information et de l'ergonomie de l'approche. Pour cela, nous avons organisé ce manuscrit en quatre chapitres.

La revue de questions dresse un état de l'art des différents aspects de la problématique de l'accès à l'information. Celui-ci s'articule autour des différents éléments de réponse à la question : « quels sont les différentes approches existantes et les différents outils correspondants qui permettent d'améliorer l'accès à l'information ? ». Cette question nous conduit tout d'abord à identifier le facteur déclenchant d'une démarche d'accès à l'information : le problème d'information. Cette notion, mise notamment en exergue par [Marchionini 95], permet de distinguer plusieurs niveaux de démarches d'accès à l'information qu'il convient de traiter différemment. Nous l'avons utilisé pour définir un cadre méthodologique pour notre approche de résolution de cette problématique.

Nous nous intéressons ensuite aux deux types de stratégies d'accès à l'information classiquement distinguées dans la littérature : les stratégies d'interrogation et de parcours. Pour chacune d'elles, une description de principe est donnée, et les outils informatiques permettant sa mise en œuvre (i.e. les moteurs de recherche et les systèmes hypermédias d'accès à l'information) sont décrits. Des éléments de résolution de la problématique d'accès à l'information à l'aide de ces outils sont décrites, comme l'indexation des fonds documentaires préalable à leur interrogation et les techniques de structuration des hypertextes. Il apparaît que les stratégies d'interrogation et de parcours sont pertinentes dans des contextes d'accès à l'information différents. Une identification des caractéristiques de ces contextes est alors proposée. Il apparaît qu'il n'existe pas de distinction nette entre ces deux approches, mais un passage continu de l'une à l'autre de ces stratégies, voire une alternance de leur utilisation par les individus. Cette conclusion nous conduit à examiner les différentes approches déjà menées pour alterner ou hybrider ces deux types de stratégies. Celles-ci sont décrites en parallèle avec

les outils correspondants. Une approche retenue dans laquelle peu de travaux ont été menés jusqu'ici est la formulation de requêtes par la navigation hypermédia.

Dans une troisième partie de ce chapitre, nous passons en revue les différentes approches permettant l'aide à l'interprétation d'un élément d'information accédé. Différentes disciplines, dérivées des sciences cognitives, nous paraissent relever de cette problématique. En particulier, la gestion des connaissances permet une approche par la recontextualisation de cette information dans l'activité qui l'utilise qu'il nous paraît intéressant d'étudier. L'ergonomie des logiciels, en tant que démarche visant à améliorer les conditions d'utilisation des outils d'accès à l'information, est également considérée comme pertinente pour l'aide à la compréhension de l'information. Nous notons d'une part les travaux effectués pour la distinction de profils des utilisateurs de systèmes d'accès à l'information. Ces travaux permettent notamment un « filtrage » de l'information disponible et une présentation de l'interface des outils personnalisés. D'autre part, les apports de l'ergonomie pour le processus de conception de systèmes hypermédias d'accès à l'information sont étudiés. Ces apports concernent la modélisation et l'analyse préalables des tâches des utilisateurs, les règles de conceptions ergonomiques et les méthodes d'évaluation et de validation de ces outils. Concernant les différentes approches de modélisation des tâches, [Tricot & Nanard 98] soulignent l'absence d'un cadre méthodologique tenant compte de l'activité motivant la démarche d'accès à l'information.

Nous concluons cet état de l'art en synthétisant les différents éléments liés à la problématique d'accès à l'information étudiés. Il nous semble alors pertinent d'étudier la pertinence d'une stratégie hybride d'accès à l'information fondée sur la recontextualisation de l'information dans l'activité dans laquelle s'inscrit la démarche d'accès à l'information. Une telle approche permet d'étudier les perspectives méthodologiques de modélisation du contexte de la tâche d'accès à l'information pour le processus de conception de l'outil. Cette approche est pluridisciplinaire. Elle s'appuie sur une méthode de modélisation du contexte d'accès à l'information issue de la gestion des connaissances, une identification des caractéristiques personnalisées des utilisateurs empruntée à l'ergonomie et une indexation conceptuelle du fonds documentaire relevant des sciences de l'information. D'un point de vue méthodologique, nous proposons d'étudier une approche de conception justifiée guidée par la modélisation des connaissances.

Les méthodes et les outils utilisés dans cette thèse sont définies au chapitre suivant. Il débute par la définition des concepts auxquels nous faisons référence dans ce travail.

Puis les méthodes utilisées dans deux des différentes disciplines sur lesquelles s'appuie notre approche sont détaillées. Nous décrivons la méthode de gestion des connaissances MKSM qui est utilisée lors de la phase de modélisation. Les différents travaux déjà menés pour l'opérationnalisation de cette méthode sont également décrits. Les aspects ergonomiques de la méthode de conception sont ensuite présentés. Contrairement à l'approche utilisée en gestion des connaissances, ces aspects sont issus d'une approche spécifique qui nécessite d'être préalablement mise au point. Elle reprend les étapes de modélisation des tâches, de mise au point des profils et d'évaluation de l'interface.

Les différentes technologies auxquelles nous avons recours pour le prototypage de l'outil sont ensuite abordées. En particulier, l'approche des documents virtuels personnalisables propose des solutions techniques pour la mise au point d'une interface dynamique prenant en compte le profil de l'utilisateur. Les autres aspects technologiques présentés sont spécifiques aux contraintes techniques rencontrées pour le développement de notre prototype.

Ces contraintes techniques sont notamment imposées par l'unité dans laquelle a été mené un projet-pilote présenté en fin de ce chapitre. Ce projet, sur lequel s'appuie le travail présenté, concerne le domaine de la maîtrise du risque de criticité nucléaire. Ce domaine, la modélisation des connaissances réalisée et la composition du fonds documentaire sont tour à tour exposés.

Nous concluons ce chapitre en rappelant les différentes méthodes choisies dans ce travail. L'intérêt d'une méthode de modélisation « unifiant » les méthodes empruntées aux différentes disciplines est souligné. L'utilisation de la méthode MKSM comme support de la démarche de conception logicielle est proposé.

La contribution de la thèse est exposée au chapitre suivant. Celui-ci commence par une analyse de la problématique soulevée en conclusion de l'état de l'art. D'un point de vue méthodologique, notre approche de l'accès à l'information par la recontextualisation et spécifique à l'utilisateur est décomposée en trois modèles conceptuels. Le modèle du domaine concerne le contexte de l'accès à l'information identifié par la modélisation des connaissances. Celui de l'utilisateur permet de modéliser les caractéristiques pertinentes des individus prises en compte pour l'accès à l'information. Le modèle de l'information permet de faire correspondre les éléments du fonds documentaire aux caractéristiques du besoin d'information définies par les deux autres modèles. Pour chacun de ces trois modèles, deux points sont à étudier. D'une part, ils doivent être mis au point en fonction des techniques présentées dans les chapitres précédents. D'autre part, nous soulignons dans cette

problématique que la méthodologie de modélisation permettant de les définir doit être optimisée pour permettre l'application concrète de l'approche. L'objectif est la réutilisabilité de la méthode de conception présentée. La problématique théorique concerne également la stratégie d'accès à l'information spécifique aux trois modèles conceptuels. Il s'agit d'identifier une approche d'hybridation des stratégies analytiques et de parcours permettant la sélection et l'articulation des éléments de contexte de l'activité.

La problématique opérationnelle est ensuite abordée. Elle concerne dans un premier temps la mise au point de l'outil d'accès à l'information permettant la mise en œuvre de l'approche. Par ailleurs, il s'agit d'évaluer la pertinence de l'utilisation de la méthode de gestion des connaissances pour la modélisation et la démarche de conception de l'outil.

Une troisième partie de cette problématique concerne l'apport spécifique de l'une de trois disciplines sur lesquelles se fonde notre démarche : la gestion des connaissances. Il s'agit d'introduire une discussion sur la pertinence de cette discipline par rapport à l'approche proposée dans ce travail. Cette discussion, qui est entrevue d'un point de vue méthodologique et théorique fait l'objet du chapitre suivant.

La deuxième partie de la contribution est l'étude de conception du prototype. Elle concerne la mise au point des modèles et la stratégie spécifique (baptisée « cumulative ») d'accès à l'information. Cette stratégie s'appuie sur un formulaire spécifique (le « caddie à concepts ») qui permet à l'utilisateur de sélectionner et d'articuler les éléments de son besoin d'information grâce à la navigation hypermédia dans la modélisation de son activité. L'approche utilise ensuite un système de règles pour formaliser des requêtes spécifiques au profil de l'utilisateur, à l'indexation conceptuelle du fonds documentaire et au langage de requêtes du moteur de recherche utilisé. Les solutions techniques adoptées pour la conception du prototype sont présentées.

L'analyse des tâches des différents types d'utilisateurs est ensuite menée. Elle aboutit à la spécification des différents composants de l'interface. Il apparaît notamment que notre stratégie d'accès à l'information n'est pas pertinente dans les cas extrêmes, c'est-à-dire pour les utilisateurs novices dans le domaine et experts en recherche d'information sur ce domaine. Dans ces cas, nous proposons des solutions plus appropriées. Les éléments d'interface sont analysés du point de vue de leurs qualités ergonomiques en fonction des critères présentés dans [Bastien & al. 98].

Ce n'est qu'après la spécification de l'interface que les différentes fonctionnalités du système sont synthétisées et analysées d'un point de vue dynamique. Cet ordonnancement des tâches de conception permet de prendre en compte les conclusions de l'analyse des critères ergonomiques préalablement effectuée. Une description fonctionnelle globale de l'architecture de l'outil est alors présentée.

Le cycle de vie, et notamment les problèmes liés à la maintenance de cet outil sont alors envisagés. Il ressort que les opérations de maintenance liées au fonctionnement « normal » (i.e. synchronique) du système peuvent être automatisées. Les modifications liées à la modélisation sont également discutées. L'objectif est de déterminer dans quelle mesure elles peuvent être prises en compte dans les modèles sans entraîner l'obsolescence de l'outil.

Les résultats concrets sont ensuite présentés. Le prototype spécifié n'est pas totalement achevé à la date de rédaction de ce manuscrit. Toutefois, différents éléments de ce prototype ont déjà été réalisés pour des démonstrateurs et nous escomptons des développements rapides de ce prototype.

La discussion fait l'objet du dernier chapitre de cette thèse. Dans une première partie, la démarche pluridisciplinaire d'accès à l'information par la recontextualisation dans l'activité des utilisateurs est discutée. Dans un premier temps, l'intérêt et les limites de cette approche sont présentés dans le cadre général de la mise en relation des trois disciplines impliquées.

Puis, nous passons en revue, d'un point de vue ergonomique, l'utilité des formalismes diagrammatiques pour l'aide à la formulation du besoin d'information. Pour chacun d'eux, nous indiquons les catégories de caractéristiques du besoin d'information qu'ils permettent d'identifier. En particulier, nous détaillons les aspects ergonomiques de la sélection des caractéristiques du besoin d'information par la navigation hypermédia dans les diagrammes de connaissances issus de la modélisation.

Nous rappelons ensuite l'apport de l'ingénierie de l'information pour notre approche. Il s'agit de discuter le paradigme d'indexation conceptuelle. Celui-ci permet de relier la modélisation des connaissances et les sources documentaires.

Enfin, la dernière section de cette partie concerne la pertinence de la méthode de gestion des connaissances MKSM dans le travail présenté. La pertinence de l'utilisation de la méthode de gestion des connaissances MKSM pour la démarche transdisciplinaire de conception de l'outil d'accès à l'information est évaluée. Il s'agit de définir l'utilisabilité de chacun des formalismes diagrammatiques pour les différentes étapes de la mise au point des modèles

conceptuels et de l'interface. Nous abordons ensuite le positionnement de la méthode MKSM dans un cadre plus général. Nous discutons les paradigmes systémique et sémiotique sur lesquels cette méthode est bâtie. Cette discussion nous permet d'aboutir à une représentation plus pragmatique du positionnement des modèles. Celle-ci peut être utilisée pour présenter la structure hypermédia du modèle du domaine à l'utilisateur du système d'accès à l'information réalisé. Des éléments de réflexion plus théoriques sont ensuite proposés afin d'évaluer la pertinence de l'approche récursive qui consiste à réinjecter dans l'activité d'une organisation la modélisation de cette activité obtenue en gestion des connaissances. Il apparaît que le succès de cette démarche est lié au degré de corrélation de l'application visée par la gestion des connaissances et du domaine d'activité modélisé.

La seconde partie de ce chapitre récapitule les conclusions tirées dans cette discussion. Elle permet d'envisager des perspectives d'évolution techniques de l'outil. Dans un premier temps, le détail de la phase d'évaluation et de validation est présenté. Puis des perspectives d'évolution de l'outil réalisé sont proposées. L'intégration des règles de formalisation des requêtes et de comportement de l'interface et l'accès à l'information non textuelle sont précisés. Enfin, une perspective d'extension de l'approche d'accès à l'information pour la génération de requêtes en langage naturel est évoquée.

Plus généralement, ces travaux ont pour objectif d'envisager les différents aspects méthodologiques de l'apport de la modélisation multi-points-de-vue utilisée en gestion des connaissances pour l'accès à l'information. L'approche multidisciplinaire de la stratégie d'accès à l'information proposée et l'approche transdisciplinaire de la démarche de conception nous ont permis de proposer des perspectives d'opérationnalisation de cette modélisation souvent pertinentes, parfois inattendues. La richesse de la modélisation obtenue est à notre sens l'un des principaux facteurs de succès de cette approche.

Nous pensons que cette modélisation peut également être utilisée avec succès dans d'autres disciplines, moyennant les restrictions énoncées au chapitre 5. Pour cela, il est toutefois nécessaire que l'application envisagée soit compatible avec le caractère qualitatif de la modélisation. Celui-ci rend à notre sens l'opérateur humain indispensable dans le processus de transformation des modèles obtenus¹⁵⁷. En cela, il convient de distinguer les perspectives d'opérationnalisation de celles d'automatisation. Nous soulignons d'ailleurs à la fin du chapitre 5 que des applications de la méthode MKSM à d'autres disciplines sont à l'étude.

¹⁵⁷ Excepté dans les applications issues de l'Intelligence artificielle, telles que les systèmes experts.

Concernant l'application de la modélisation issue de la gestion des connaissances à d'autres disciplines, les sciences de l'information constituent la discipline la plus directement visée par cette application (cf. par exemple [Dupoirier & Ermine 00]). Nous avons d'ailleurs souligné au chapitre 3 le fait que la gestion des connaissances est un sous-domaine des sciences de l'information. Nous avons également montré que dans ce domaine, les différents formalismes de la méthode peuvent être utilisés pour structurer la processus de conception, à l'instar d'une méthode de génie logiciel.

Dans la perspective du prolongement de ce travail, il nous semble intéressant d'étudier l'utilité de la gestion des connaissances dans le cadre plus global du cycle de vie du système d'information de l'organisation modélisée. En effet, attendu que le système d'information possède trois fonctions qui sont la production, le stockage et l'accès à l'information, nous n'avons exploré ici que l'une de ces fonctions. La fonction de stockage est par définition peu intéressante. Elle se résume à la définition initiale des outils adéquats. La fonction de production de l'information présente des caractéristiques qu'il peut être intéressant d'explorer.

En se fondant sur des formalismes flots de données, de type SADT, on pourra notamment s'intéresser à l'activité de production documentaire en tant que réutilisation de l'information existante. La pertinence de la gestion des connaissances en tant qu'interface entre la documentation existante et l'information à rédiger pourra être considérée. Le statut des formalismes diagrammatique sera à reconsidérer spécifiquement à cet objectif. Pour cela, une approche d'intégration de la modélisation de l'activité et des outils de production documentaire existants devra être étudiée. L'adaptation de notre approche à la conception d'un « système d'aide à la rédaction technique » pourra être envisagée. Une réflexion sur l'impact de l'approche diagrammatique sur le mode d'interaction à définir pour la rédaction nous paraît prometteuse.

Bibliographie

- [Bastien & al. 98] : Bastien, J. M. C., Leulier C., Scapin D-L, "*L'Ergonomie des sites web*", in Créer et Maintenir un service web, J-C. Le Moal & B. Hidoine Eds., ADBS, Paris, pp. 111-173, 1998.
- [Dupoirier & Ermine 00] : "*Gestion des documents et gestion des connaissances*", G. Dupoirier et J-L. Ermine (Coordonnateurs), Document Numérique, Hermès, vol. 3, n° spécial 3-4, Paris, 1999.
- [Marchionini 95] : Marchionini G., "*Information seeking in electronic environments*", New York, Cambridge University Press, 1995.
- [Tricot & Nanard 98] : Tricot A., Nanard J., "*Un point sur la modélisation des tâches de recherche d'information dans le domaine des hypermédias*", in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 35-56, 1998.

Annexes

Annexe 1. Liste des références bibliographiques

- [ACM 92] : Association for Computer Machinery, Special Interest Group on Computer-Human Interaction, *"Curricula for Human-Computer Interaction"*, ISBN 0-89791-474-0, disponible en ligne : <http://www.acm.org/sigchi/cdg/cdg2.html>, 1992.
- [Addis 96] : Addis, T. R., Addis, J. J. T., *"A Functional Schema Interpreter as an aid to Knowledge System Design"*, in proceedings of the Thinking with Diagrams colloquium, IEE Computing and Control Division, Professional Group C4 (Artificial Intelligence), Digest No: 96/010, pp. 13/1-13/10, London, 18 janvier 1996.
- [Adobe 99] : Adobe Developer Technologies, *"Highlight File Format"*, Technical Note #5172, Adobe Systems Inc., date de dernière révision de la version utilisée : 10 nov. 1999, disponible en ligne : <http://partners.adobe.com/asn/developer/technotes/TN5172.pdf>.
- [Agosti & al. 91] : Agosti M., Colotti R., Gradenigo G., *"A two-level hypertext retrieval model for legal data"*, in proceedings of the ACM Hypertext'91 Conference, pp. 316-325, 1991.
- [Agosti & Smeaton 96] : *"Information Retrieval and Hypertext"*, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, 1996.
- [Ahlberg & al. 92] : Ahlberg C., Williamson C., Shneiderman B., *"Dynamic Queries for Information Exploration: An Implementation and Evaluation"*, in proceedings of the ACM CHI'92 Conference on Computer-Human Interaction, pp. 619-626, Mai 1992.
- [Ahlswede 88] : Ahlswede T., *"Automatic construction of phrasal thesaurus for an information retrieval system form a machine readable dictionnary"*, RIAO 88, vol. 1, pp. 597-608, 1988.
- [Alix & Roussel 01] : Alix C., Roussel C., *"La vie est mouvante, c'est cette complexité qu'il faut représenter"*, interview de T. Nelson, in Week-end Rencontre, Libération, pp. 36-37, édition du 26 & 27 mai 2001.
- [Allen 97] : Allen R. B., *"Mental Models and User Models"*, Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu Eds., Elsevier Science, 1997, pp.49-63.
- [Assadi & Gros 99] : Assadi H., Gros C., *"Ingénierie des connaissances pour les systèmes de consultation de documentation technique"*, Document Numérique, Hermès, vol. 3, n° 3-4, pp. 15-35, 1999.
- [Asse 85] : Asse A., *"Aide au diagnostique industriel par des méthodes basées sur la théorie des sous-ensembles flous"*, thèse de doctorat de troisième cycle, Université de Valenciennes, avril 1985.
- [Balpe & al. 96] : Balpe J-P., Lelu A., Papy F., Saleh I., *"Techniques avancées pour l'hypertexte"*, Hermès, Paris 1996.
- [Barker 97] : Barker A. L., *"DataStar Web: living up to the hype? An evaluation of the interface and search system"*, in proceedings of the Online information'97 Conference, pp. 213-222, 1997.
- [Barthes 64] : Barthes R., *"Éléments de sémiologie"*, 1964.
- [Bastien & al. 98] : Bastien, J. M. C., Leulier C., Scapin D-L., *"L'Ergonomie des sites web"*, in Créer et Maintenir un service web, J-C. Le Moal & B. Hidoine Eds., ADBS, Paris, pp. 111-173, 1998.

- [Bastien & Scapin 93] : Bastien, J. M. C., Scapin D-L, "*Ergonomic Criteria for the Evaluation of Human-Computer Interfaces*", Rapport technique INRIA n° 156, Juin 1993.
- [Bates 89] : Bates M., "*The design of browsing and berrypicking techniques for the on-line search interface*", Online Review, 13(5), pp. 407-431, 1989.
- [Bateson 87] : Bateson A. G., Alexander R. A., Murphy M. D., "*Cognitive processing differences between novice and expert computer programmers*", International Journal of Man-Machine Studies, vol. 26, pp. 649-660, 1987.
- [Belkin & al. 93] : Belkin N. J., Marchetti P. G., Cool Chen, "*BRAQUE: Design of an interface to support user interaction in information retrieval*", Information Processing and Management, vol. 29, n° 3, pp. 325-344, 1993.
- [Belkin 80] : Belkin N. J., "*Anomalous states of knowledge as a basis for Information retrieval*", Canadian Journal of Information Science, n°5, pp. 133-143, 1980.
- [Bernstein 90] : Bernstein M., "*Hypertext and technical writing*", in proceedings of the ECHT'90 Course, Versailles, Novembre 1990.
- [Bernstein 93] : Bernstein M., "*Enactment in information farming*", in proceedings of HyperText'93 Conference, ACM Press, pp. 242-249, 1993.
- [Blanger 98] : Blanger J-P., Notes personnelles prises lors de la formation au logiciel Acrobat, BVA MYFRA, octobre 1998.
- [Booch 92] : Booch G., "*Conception orientée objets et applications*", Addison-Wesley, 1992.
- [Brusilovsky 96] : Brusilovsky P., "*Methods and techniques of adaptive hypermedia*", User Modeling and User-Adapted Interaction, Special issue on adaptive hypertext and hypermedia, 6 (2-3), pp. 87-129, 1996, disponible en ligne : <http://citeseer.nj.nec.com/cache/papers2/cs/440/http:zSzzSzwww.contrib.andrew.cmu.edu/Sz~plbzSzUMUAI.pdf/brusilovsky96methods.pdf>
- [Buckingham Shum 96] : Buckingham Shum S., "*Design Argumentation as Design Rationale*", in The Encyclopaedia of Computer Science and Technology, Kent A. and William J. G. (Eds), Marcel Dekker Inc., vol. 35, Suppl. 20, pp. 95-128, New York, disponible en ligne : <ftp://kmi-ftp.open.ac.uk/pub/simonb/ecst96.pdf>, 1996.
- [Buckland 91] : Buckland M., "*Information as thing*", journal of the American Society of Information Science, vol. 42, N° 5, pp. 351-360, disponible en ligne : <http://www.sims.berkeley.edu/~buckland/thing.html>, 1991.
- [Bush 45] : Bush V., "*As we may think*", Atlantic Monthly, 176 (1), pp. 101-108, 1945.
- [Chaillot & Ermine 97] : Chaillot M., Ermine J-L., "*Le livre de connaissances électronique*", Document numérique, 1(1), 75-98, 1997.
- [Chang & Rice 93] : Chang S-J., Rice R-E., "*Browsing: a multidimensional framework*", Annual review of information science and technology, Vol.28, Learned Inf., Medford, NJ, USA; xv, 1993, pp. 231-76.
- [Charlot & al. 00] : Charlot J-M., Chaillot M., Médini L., "*Information Engineering based on knowledge modelling: a case study in nuclear criticality safety*", International Seminar on Integrated Information Systems, IAEA-SR-212/46, Vienne (Austria), april, 10-14, 2000.
- [Charreton & Ermine 96] : Charreton B., Ermine J-L., "*From knowledge specification to executable specification*", in proceedings of KEML'96, Knowledge Engineering and Modelling Languages, Paris, 1996.

- [Charron & Fluhr 99] : Charron J., Fluhr C., *"BeFor (Beyond Forms). Un modèle de représentation du Web invisible"*, Actes de la conférence H²PTM'99, Hermès, pp. 27-41, Paris, 1999.
- [Chiamarella & Kheirbek 96] : Chiamarella Y., Kheirbek A., *"An intergrated model for hypermedia and information retrieval"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 139-178, 1996.
- [Claverie 97] : Claverie B., Le Blanc B., Corsini M., *"Étude des processus cognitifs privilégiés dans l'exploitation des bases de connaissances pour une navigation dynamique liée à l'utilisateur"*, sujet de thèse proposé à l'université de Bordeaux, 1997.
- [Cooper & al. 1990] : Cooper R., Mukai K., Perry J., *"Situation Theory and its Applications"*, vol. 1, CSLI, 1990.
- [Crampes & Ranwez 00] : Crampes M., Ranwez S., *"Ontology Supported and Ontology-Driven Conceptual Navigation on the World Wide Web"*, in proceedings of the ACM HyperText'2000 Conference, pp. 191-199, 2000.
- [Crampes 99] : Crampes M., *Introduction à l'atelier sur les Documents Virtuels Personnalisables*, in Interaction Homme-Machine, Actes de l'Atelier Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, Novembre 1999.
- [de Rosnay 75] : de Rosnay J., *"Le macroscopie : vers une vision globale"*, Paris: Seuil, 1975.
- [de Saussure 16] : de Saussure F., *"Cours de linguistique générale"*, 1916.
- [de Vriès & de Jong 97] : de Vriès E., de Jong T., *"Using information systems while performing complex tasks: an example from architectural design"*, International Journal of Man-Machine Studies, vol. 46, n° 1, pp. 31-54, 1997.
- [Dechilly & al. 99] : Dechilly T., Auffret G., Brunie V., Bachimont B., *"Représentation et structuration des connaissances pour les bibliothèques audiovisuelles"*, Document Numérique, vol. 3, n° 3-4, Hermès, pp. 195-214, Paris, 1999.
- [Dervin & Nilan 86] : Dervin B., Nilan M., *"Information needs and uses"*, Annual Review of Information Science and Technology, M. Williams (Ed.), White Plains, New York: Knowledge Industries, vol. 21, No. 3, pp. 33, 1986.
- [Dinet & al. 98] : Dinet J., Rouet J-F, Passerault J-M., *"Les « Nouveaux outils » de recherche documentaire sont-ils compatibles avec les stratégies cognitives des élèves ?"*, in actes du 4^{ème} Colloque Hypermédias et Apprentissages, J-F Rouet et B. de la Passardière Eds., Poitiers, pp. 149-161, octobre 1998.
- [DISTNB 98] : Direction de l'information scientifique, des technologies nouvelles et des bibliothèques (DISTNB), Département de l'information spécialisée, Définition de l'IST disponible en ligne : <http://www.mesr.fr/icst/ist/index.htm>, 1998.
- [Dunlop & van Rijsbergen 93] : Dunlop M., Van Rijsbergen C. J., *"Hypermedia and free text retrieval"*, Information Processing and Management, vol 29, n° 3, pp. 287-298, 1993.
- [Dupoirier & Ermine 00] : *"Gestion des documents et gestion des connaissances"*, G. Dupoirier et J-L. Ermine (Coordonnateurs), Document Numérique, Hermès, vol. 3, n° spécial 3-4, Paris, 1999.
- [Eco 88] : Eco, U., *"Le signe"*, Livre de poche, Essais, 1988.
- [Engelbart 62] : Engelbart D. C., *"Augmenting human intellect: A conceptual framework"*,

- Technical Report AF 49(638)-1024, Stanford Research Institute, Octobre 1962.
- [Erdmann & Studer 00] : Erdmann M., Studer R., *"How to Structure and Access XML Documents With Ontologies"*, Data and Knowledge Engineering, special issue on Intelligent Information Integration, 2000.
- [Ermine & al. 96] : Ermine J-L., M. Chaillot, P. Bignon, B. Charreton, D. Malavieille: *"MKSM, a method for knowledge systems management"*. in Knowledge Management, Organization, Competence and Methodology, Advances in Knowledge Management, Volume 1, Jos F. Schreinmakers Ed., pp. 288-302, Ergon Verlag, 1996.
- [Ermine 89] : Ermine J-L, *"Systèmes experts, théorie et pratique"*, Collection Tec et Doc, Lavoisier Ed, Paris, 1989.
- [Ermine 96] : Ermine J-L., *"Les systèmes de connaissances"*, Hermès, 1996.
- [Ermine 97] : Ermine J-L., *"MKSM, méthode pour la gestion des connaissances"*, polycopié de cours n° 9, DEA SETI, CEA Saclay, 1997.
- [Farenc & Palanque 99] : Farenc C., Palanque P., *"Exploitation des notations de Design Rationale pour une conception justifiée des applications interactives"*, in Interaction Homme-Machine, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Cépaduès-Éditions, Toulouse, France, vol. 1, pp. 33-40, 1999.
- [Farenc 00] : Farenc C., *"ERGOVAL : une méthode de structuration des règles ergonomiques permettant l'évaluation automatique d'interfaces graphiques"*, Thèse de doctorat, Chapitre 1 – Les méthodes d'évaluation, version du 11 décembre 2000, disponible en ligne : <http://lis.univ-tlse1.fr/farenc/nouvelle.htm>.
- [Felber 87] : Felber H., *"Manuel de terminologie"*, Paris, UNESCO, 1987.
- [Finke 89] : Finke R. A., *"Principles of mental imagery"*, MIT Press, Cambridge, Mass., 1989.
- [Fisher & al. 96] : Fisher G., Lemke A. C., Mc Call R., Morch A. I., *"Making Argumentation Serve Design"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp.393-419, réédité dans Moran T. P. & Carroll J. M. (Eds.), Design Rationale: Concepts, Techniques and Use, Lawrence Erlbaum Associates, pp. 267-293, 1996.
- [Fluhr & al. 97] : Fluhr C., Schmit D., Ortet P., Elkateb F., Gurtner K., *"SPIRIT-W3: a distributed crosslingual indexing and search engine"*, in proceedings of the 7th Annual Conference of the Internet Society INET'97, the Internet, the Global Frontier, Juin 97.
- [Foucault 99] : Foucault B., *"Analyse des stratégies de navigation sur le Web : expériences sur les modes d'accès à l'information"*, in Interaction Homme-Machine, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Cépaduès-Éditions, Toulouse, France, vol. 2, pp. 45-48, 1999.
- [Fucella & al. 98] : Fucella j., Pizzolato J., Franks J., *"Web site user centered design: technique for gathering requirements and tasks"*, Internetworking, 1 (1), disponible en ligne : http://www.sandia.gov/itg/newsletter/june98/web_design.html, 1998.
- [Fucella & Pizzolato 98] : Fucella j., Pizzolato J., *"Creating Web site design based on user expectations and feedback"*, Internetworking, 1 (1), disponible en ligne : http://www.sandia.gov/itg/newsletter/june98/user_requirements.html, 1998.
- [Ganascia 93] : Ganascia J-G., *"Les Sciences Cognitives"*, Collection Que Sais-Je ?, Dunod, Paris, 1993.

- [Garlatti & Iksal 99] : Garlatti S., Iksal S., "*Documents virtuels personnalisables pour les systèmes d'information*", in Interaction Homme-Machine, Actes de l'Atelier sur les Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, pp. 22-26, Novembre 1999.
- [Garzotto & al. 91] : Garzotto F., Mainetti L., Paolini P., "*HDM: a model for the design of hypertext applications*", in proceedings of the ACM Hypertext'91 Conference, pp. 313-328, décembre 1991.
- [Garzotto & al. 93] : Garzotto F., Mainetti L., Paolini P., "*HDM: a model based approach to hypertext application design*", in proceedings of the ACM Transactions on Information Systems, vol. 11, n° 1, pp. 1-26, 1993.
- [Garzotto & al. 95] : Garzotto F., Paolini P., Mainetti L., "*Hypermedia Design, Analysis and Evaluation Issues*", in Communications of the ACM, n° 38, vol. 8, pp. 74-86, Août 1995.
- [Garzotto & al. 99] : Garzotto F., Matera M., Paolini P., "*Abstract Tasks: a Tool for the inspection of Web Sites and Off-line Hypermedia*", in proceedings of the tenth ACM Conference on Hypertext and hypermedia, pp. 157-163, 1999.
- [Glasgow & al. 95] : Glasgow J., Narayanan, H., Chandrasekaran B., "*Diagrammatic Reasoning, Cognitive and Computational Perspectives*", The MIT Press, Massachusetts Institute of Technology, Cambridge, Massachusetts, London, England, 1995.
- [Golovchinsky & Chignell 93] : Golovchinsky G., Chignell M., "*Queries-R-Links: Graphical Markup for Text Navigation*", in proceedings of the ACM INTERCHI'93, Conference on Human Factors in Computing Systems, pp. 454-460, 1993.
- [Golovchinsky 97] : Golovchinsky, G., "*Queries? Links? Is there a difference?*", in proceedings of the ACM CHI'97 Conference on Human Factors in Computing Systems, vol. 1, pp. 407-414, 1997.
- [Guarino & al. 99] : Guarino N., Masolo C., Vetere G., "*OntoSeek: Content-Based Access to the Web*", in IEEE Intelligent Systems, pp. 70-80, Mai-Juin 1999.
- [Halasz & Schwartz 94] : Halasz F. G., Schwartz M., "*The Dexter hypertext reference model*", Communication of the ACM, vol. 37, n° 2, pp.30-39, 1994.
- [Hammond & Allinson 89] : Hammond N., Allinson L., "*Extending Hypermedia for Learning: An Investigation of Access and Guidance Tools*", in Proceedings of the HCI'89 Conference on People and Computers, pp. 293-304, 1989.
- [Hardman & al. 94] : Hardman L., Bulterman D., Van Rossum G., "*The Amsterdam hypermedia model*", Communication of the ACM, vol. 37, n° 2, pp. 50-63, 1994.
- [Hearst & Karadi 97] : Hearst M. A., Karadi C., "*Cat-a-Cone: An Interactive Interface for Specifying Searches and Viewing Retrieval using Large Category Hierarchy*", in proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, p.246-255, 1997.
- [Heflin & al. 99] : Heflin, J., Hendler J., Luke S., "*SHOE: a Knowledge Representation Language for Internet Applications*", Technical Report CS-TR-4078 (UMIACS TR-99-71), Dept of Computer Science, University of Maryland at College Park, 1999, disponible en ligne : <http://www.cs.umd.edu/projects/plus/SHOE/pubs/techrpt99.pdf>
- [Huibers & al. 96] : Huibers T., Ounis I., Chevallet J-P., "*Conceptual Graph Aboutness*", in proceedings of the 4th International Conference on Conceptual Structures, ICCS'96, Lecture Notes in Artificial Intelligence, P. W. Eklund, G. ELLIS, G. Mann Eds., pp. 130-145, Sydney, 1996.

- [Hunter & al. 01] : Hunter D., Cagle C., Gibbons D., Ozu N., Pinnock J., Spencer P., *"Initiation à XML. Avec trois études de cas détaillées"*, Traduit par F. Lemainque, L. Adam, C. Raspaud, Eyrolles, Paris 2001.
- [IGL 82] : IGL, France : *"Introduction à SADT"*, Manuel, 1982.
- [Ingwersen 94] : Ingwersen P., *"Polyrepresentation of information needs and semantics entities: elements of a cognitive theory for information retrieval interaction"*, in Proceedings of the seventh Annual International ACM-SIGIR, Dublin, Ireland, 1994.
- [Iwayama & Tokunaga 95] : Iwayama M., Tokunaga T., *"Cluster-Based Text Categorization: A Comparison of Category Search Strategies Text Categorization"*, in proceedings of the 8th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval p.273-280, 1995.
- [John & Kieras 94] : John, B. E. , Kieras, D. E., *"The GOMS family of analysis techniques : tools for design and evaluation"*, Carnegie Mellon University, Pittsburg, 1994.
- [Johnson-Laird 83] : Johnson-Laird P. N., *"Mental models"*, Harvard University Press, Cambridge, Mass., 1983.
- [Jolion 01] : *"L'indexation"*, numéro spécial de la revue Document Numérique, sous la direction de J-M Jolion, Document Numérique, Vol 4, n°1-2/2000, Hermès Science Publication, 2001.
- [Jouis, 93] : Jouis C., *"Contributions à la conceptualisation et à la modélisation des connaissances à partir d'une analyse linguistique de textes"*, Thèse de doctorat de l'École des Hautes Études en Sciences Sociales, 1993.
- [Kass & Finin 88] : Kass T., Finin T., *"The Need for User Models in Generating Expert System Explanations"*, in International Journal of Expert Systems Research and Applications, vol. 1., pp. 345-375, 1988.
- [Kintsch & Van Dijk 78] : Kintsch W., Van Dijk T. A., *"Towards a Model of Text Comprehension and Production"*, Psychological Review, vol. 85, n° 5, pp. 363-394, 1978.
- [Kolmayer 98] : Kolmayer E., *"Démarche d'interrogation documentaire et navigation"*, in Quatrième Colloque Hypermedias et Apprentissages, Poitiers, pp. 89-96, 15-17 octobre 1998.
- [Kolski 97] : Kolski C., *"Interfaces homme-machine, application aux systèmes industriels complexes"*, 2^{ème} édition revue et augmentée, Hermès, Paris, 1997.
- [Kosslyn 80] : Kosslyn S. M., *"Image and Mind"*, Harvard University Press, Cambridge, Mass., 1980.
- [Kristen 98] : Kristen G., transparents issus de la présentation de la méthode KISS, organisée au CEA de Saclay en avril 1998.
- [Kuntz & Rittel 70] : Kuntz W., Rittel H. W. J., *"Issues as elements of information systems"*, Working paper No. 131, Berkeley, UCB, Center for Planning and Development Research, 1970.
- [Lainé-Cruzel & al. 96] : Lainé-Cruzel S., Lafouge T., Lardy J-P. & Ben Abdallah N., *"Improving information retrieval by combining user profile and document segmentation"*, Information Processing and Management, Elsevier Science Ltd, Vol. 32, No. 3, pp. 305-315, 1996.
- [Lainé-Cruzel 94] : Lainé-Cruzel S., *"Vers de nouveaux systèmes d'information prenant en*

- compte le profil des utilisateurs*", Documentaliste – Sciences de l'information, vol. 31, n° 3, pp. 143-147, 1994.
- [Le Loarer 1994] : Le Loarer P., *"Indexation automatique, recherche d'information et évaluation"*, Cours INRIA, Le traitement électronique du document, Aix-en-Provence, pp. 149-201, octobre 1994.
- [Le Moigne 77] : Le Moigne, J-L., *"La théorie du système général, Théorie de la modélisation"*, Paris, PUF, 1977, 3^{ème} édition mise à jour, 1990.
- [Le Moigne 90] : Le Moigne, J-L., *"La modélisation des systèmes complexes"*, Dunod, 1990.
- [Lee & Lai 91] : Lee J., Lai K., *"What's in Design Rationale?"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp.251-280, réédité dans Moran T. P. & Carroll J. M. (Eds.), *Design Rationale: Concepts, Techniques and Use*, Lawrence Erlbaum Associates, pp. 125-154, 1996.
- [Leloup 98] Leloup, C., *"Moteurs d'indexation et de recherche, environnements client-serveur, internet et intranet"*, Eyrolles, Paris, 1998
- [Lelu & al. 99] : Lelu A., Hallab M., Rhissassi H., Papy F., Bouyahi S., Bouhaï N., He H., Qi C., Saleh I., *"Projet NeuroWeb : un moteur de recherche multilingue et cartographique"*, actes de la 5^{ème} conférence internationale Hypertextes, Hypermédias et Internet, H²PTM'99, réalisations, outils et méthodes, Hermès, Paris, pp. 43-56, 1999.
- [Lelu 93] : Lelu A., *"Modèles neuronaux pour l'analyse de données documentaires et textuelles – organiser de très grands tableaux de données qualitatives en pôles et zones d'influence"*, Thèse de l'université Paris 6, 1993.
- [Lesk 89] : Lesk M., *"What to do when there's too much information"*, in proceedings of the ACM Hypertext'89 conference, pp. 305-318.
- [Leulier & al. 98] : Leulier C., Bastien J. M. C., Scapin D-L., *"Compilation of ergonomic guidelines for the design and evaluation of websites"*, in Commerce and Interaction Report, INRIA, Rocquencourt, France, 1998.
- [Lippold & al. 94] : Lippold B., Pomian J., Henry, J. Y., Elsensohn O., *"AVIS : une méthode d'analyse de la cohérence des documents"*, in proceedings of the European Safety and Reliability Conference, pp. 888-899, La Baule, 1994.
- [Lucarella & al. 93] : Lucarella D., Parisotto S., Zanzi A., *"MORE: Multimedia Object Retrieval Environment"*, in proceedings of the ACM Hypertext'93, pp. 39-50, 1996.
- [Lucarella & Zanzi 96a] : Lucarella D., Zanzi A., *"A Visual Retrieval Environment for Hypermedia Information Systems"*, in proceedings of the ACM Transaction on Information Systems, vol. 14, n° 1, pp. 3-29, 1996.
- [Lucarella & Zanzi 96b] : Lucarella D., Zanzi A., *"Information modelling and retrieval in hypermedia systems"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 121-138, 1996.
- [Lucarella 90] : Lucarella D., *"A Model for Hypertext-Based Information Retrieval. Formal Models and Query Languages"*, in Proceedings of the ECHT'90 European Conference on HyperTexts, pp. 81-94, 1990.
- [Mackay 60] : Mackay D. M., *"What makes the question"*, The listener, n°63, pp. 789-790, 1960.
- [Mackay & al. 98] : Mackay W.E., Fayard A-L., Frobert L., Médini L., *"Reinventing the*

- familiar: Exploring an augmented reality design space for air traffic control*", in Proceedings of the ACM CHI'98, Conference on Human Factors and Computing Systems, ACM Press, pp. 558 - 565, Los Angeles, CA, 1998.
- [Mackay 88] : Mackay, W. E., *"More than Just a Communication System: Diversity in the Use of Electronic Mail"*, Proceedings of CSCW'88: Conference on Computer-Supported Cooperative Work, Portland, Oregon: ACM, 1988.
- [Mackinlay & al. 95] Mackinlay J. D., Zellweger P. T., Chignell M., Furnas G., Salton G., *"Browsing vs. Search: Can We Find a Synergy?"*, Panels, in proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems, ACM Press, vol 2, pp. 179-180, 1995.
- [Malavieille & al. 95] : Malavieille D., Ermine J-L., Ct Ribot, Lions C., Lt Condamin, *"Étude et réalisation d'un système à base de connaissances d'aide à la décision en cas d'intervention sur des accidents majeurs"*, in proceedings of IA'95, Montpellier, 1995.
- [Marchionini & Shneiderman 88] : Marchionini G., Shneiderman B., *"Finding facts vs. browsing knowledge in hypertexts systems"*, IEEE computer, vol. 21, N° 1, pp. 70-80, 1988.
- [Marchionini 89] : Marchionini G., *"Information seeking in electronic encyclopaedia"*, Machine-Mediated Learning, vol. 3, N° 3, pp. 211-226, 1989.
- [Marchionini 95] : Marchionini G., *"Information seeking in electronic environments"*, New York, Cambridge University Press, 1995.
- [Martin & Eklund 00] : Martin P., Eklund P., *"Knowledge Retrieval and the Word Wide Web"*, in IEEE Intelligent Systems, Special Issue on Knowledge Management and the Internet, Mai-Juin 2000.
- [Masand et al., 1992] : Masand B., Linoff G., Waltz D., *"Classifying news stories using memory based reasoning"*, in proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 59-65, 1992.
- [Masui & al. 96] : Masui T., Minakuchi M., Borden G. R., Kashiwagi K., *"WING: A Multiple-View Smooth Information Retrieval System"*, in proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, System Demonstrations, p. 338, 1996.
- [Mayer 97] : Mayer. R. E., *"From Novice to Expert"*, Handbook of Human-Computer Interaction, M. Helander, T. K. Landhauer, P. Prabhu eds., Elsevier Science B.V., pp.781-795, 1997.
- [Mayhew 92] : Mayhew, D. J., *"Principles and Guidelines in Software User Interface Design"*, Englewood Cliffs, NJ: Prentice-Hall, 1992.
- [Mc Call 79] : Mc Call R. *"On the structure and use of issues systems in design"*, unpublished doctoral dissertation, Berkeley, UCB, 1979.
- [Mc Knight & al. 91] : Mc Knight, C., Dillon A., Richardson J., *"Hypertext in context"*, Cambridge University Press, Cambridge, 1991.
- [Mc Lean & al. 91] : Mc Lean A., Young R. M. Bellotti V., Moran T., *"Questions, Options, and Criteria: Elements of Design Space Analysis"*, Human-Computer Interaction, special issue on Design Rationale, 6 (3-4), pp. 201-250, réédité dans Moran T. P. & Carroll J. M. (Eds.), Design Rationale: Concepts, Techniques and Use, Lawrence Erlbaum Associates, pp. 75-124, 1996.

- [Médini 97] : Médini L., *"Conception et Réalisation d'un Livre de Connaissances Électronique sous forme de Document Hypermédia"*, Rapport de projet de DEA, Orsay et INSTN, Sept. 1997.
- [Médini & al. 99] : Médini L., J. Anno, M. Chaillot, V. Rouyer, D. L. Scapin, *"Towards a knowledge oriented system for information management in nuclear safety"*, Proceedings of the Sixth International Conference on Nuclear Criticality-Safety, Versailles, Sept. 1999, 1451-1461.
- [Médini 99] : Médini L., *"Utilisation de la modélisation des connaissances pour la conception d'un référentiel hypermédia d'accès à l'information scientifique et technique"*, Rencontres Doctorales, Actes de la conférence IHM'99, vol. 2, pp. 174-177, Montpellier, Nov. 1999.
- [Médini & al. 00a] : Médini L., Bignon Ph., Charlot J-M., Chaillot M., Ferret O., Malavieille D., Penel D., Scapin D-L., Van Craeynest J-M, *"Expression du besoin d'information et indexation conceptuelle"*, Document Numérique, Hermès, vol. 4, n° 1-2, pp. 85-108, 2000.
- [Médini & al. 00b] : Médini L., Charlot J-M., Chaillot M., Scapin D-L.: *"Concevoir un Livre de Connaissances Électronique multi-utilisateurs : une étude de cas appliquée à la sûreté-criticité nucléaire"*, Actes de la conférence Nîmes'TIC 2000, 11-13 septembre 2000, Nîmes, pp. 389-399.
- [Médini & al. 01] : Médini L., Chaillot M., Charlot J-M, *"Designing an Electronic Knowledge Book: How Advanced Knowledge Management Methods Can Help Information Modeling"*, Technical Communications, Special Issue on Information Modeling, sous presse (enfin !).
- [Michard 99] : Michard A., *"XML, Langage et applications"*, Eyrolles, Paris, 1999.
- [Millerat & al. 96]: Millerat P., Ermine J-L., Chaillot M., *"Knowledge management for modelling nuclear power plants control in incidental and accidental states"*, Proceedings of the CESA'96 IMACS Multiconference, Symposium on Modelling, Analysis and Simulation, Vol. 2, pp.982-987, Lille, 1996.
- [Moran & Carroll 96] : Moran T. P., Carroll J. M. (Eds.), *"Design Rationale: Concepts, Techniques and Use"*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, ISBN : 0-8058-1566-X, 1996.
- [Motta & al. 00] : Motta E., Buckingham Shum S., Domingue J., *"Ontology-Driven Document Enrichment: Principles, Tools and Applications"*, in International Journal of Human-Computer Studies, vol. 52, n° 5, pp. 1071-1109, 2000.
- [Moulin 99] : Moulin C., Pazzaglia J-C., *"Documents pédagogiques adaptatifs dans un environnement d'apprentissage distribué"*, in Interaction Homme-Machine, Actes de l'Atelier Documents Virtuels Personnalisables, Onzièmes journées sur l'ingénierie de l'Interaction Homme-Machine, Montpellier, pp. 27-31, Novembre 1999.
- [Nanard & Nanard 89] Nanard M., Nanard J., *"MacWeb, un outil pour élaborer des documents"*, Ed. Bigre, Rennes, pp. 18-29, 1989.
- [Nanard & Nanard 91] : Nanard J., Nanard M., *"Using Structured Types to incorporate Knowledge in Hypertexts"*, in proceedings of the Hypertext'91 Conference, ACM Press, pp. 329-334, 1991.
- [Nanard & Nanard 93] : Nanard J., Nanard M., *"Should Anchors be Typed Too? An experiment with Macweb"*, in proceedings of the Hypertext'93 Conference, ACM Press, pp. 51-62, 1993.

- [Nanard & Nanard 98] : Nanard J., Nanard M., *"La conception d'hypermédias"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 15-34, 1998.
- [Nelson 67] : Nelson T. H., *"Getting it out of our system"*, in Information Retrieval: a Critical Review, G. Schechter (Ed.), Thompson Books, Washington D. C., 1967.
- [Nielsen 90] : Nielsen J., *"Hypertext & Hypermedia"*, Academic Press Inc., San Diego, Ca., 1990.
- [Nonaka 95] : Nonaka I., Takeuchi H., *"The knowledge-creating company"*, Oxford University Press, 1995.
- [Normann 86] : *"User Centered System Design: New Perspectives on Human-Computer Interaction"*, Norman D-A., Draper, S-W. (Eds), Hillsdale, NJ: Lawrence Erlbaum Associates, 1986.
- [Normann 88] : Normann D. A., *"The Psychology of Everyday Things"*, Basic Books, Inc, New York, 1988.
- [Novak & Gowin 84] : Novak J., Gowin D. B., *"Learning how to learn"*, Cambridge University Press, 1984.
- [Palmquist & Kim 98] : Palmquist R. A., Kim K-S., *"Modeling the Users of Information Systems: Some Theories and Methods"*, Electronic Resources: Use and User Behavior, Hemalata Iyer (Ed.), The Haworth Press Inc., pp. 3-25, 1998.
- [Panijel 99] : Panijel C., Unité régionale de promotion et de formation pour l'information scientifique et technique de Paris / École des Chartes, Information Scientifique et Technique, disponible en ligne : <http://www.ccr.jussieu.fr/urfist/def-ist.htm>, 1999.
- [Pardi 99] Pardi W-J., *"XML en Action"*, Microsoft Press, 1999.
- [Peeters & Charlier 95] : Peeters H., Charlier P., *"Pour une sémio-pragmatique des hypertextes multimédia : proposition théorique de catégories d'analyse pertinentes"*, disponible en ligne : <http://www.comu.ucl.ac.be/GReMS/HugoWeb/semio-hypertext.html>, 1995.
- [Perlmann & al. 90] Perlmann G., Egan D. E., Relich K., Marchionini G., Nielsen J., Shneiderman B., *"Evaluating hypermedia systems"*, in proceedings of the ACM CHI'90 Conference on Human-Computer Interaction, Panel Session, pp. 387-390, 1990.
- [Picard & al. 99] : Picard S., Ermine J-L., Scheurer B., *"Gestion des connaissances pour des grands logiciels de Calcul scientifique"*, Plate-forme AFIA, IC'99 : Ingénierie des Connaissances, pp. 171-180, juin 1999, disponible en ligne : <http://www.afia.polytechnique.fr/plate-forme/Actes/ic18.pdf>.
- [Picard 96] : Picard S., *"Conception et réalisation d'un moteur abstrait hypermédia pour la méthode MKSM"*, Orsay et INSTN, septembre 1996.
- [Pierce 78] : Pierce C. S., *"Collected Papers of Charles Sanders Pierce"*, 8 vol., Harvard University Press, Cambridge, Mass., 1978.
- [Pomian 96] : Pomian J., *"Mémoires d'entreprises, techniques et outils de la gestion du savoir"*, Paris, Sapiientia, 1996.
- [Prié 01] : Prié Y., *"Sur la piste de l'indexation conceptuelle de documents. Une approche par l'annotation"*, dans le numéro spécial de la revue Document Numérique sur l'indexation, sous la direction de J-M Jolion, Document Numérique, Vol 4, n°1-2/2000, Hermès Science

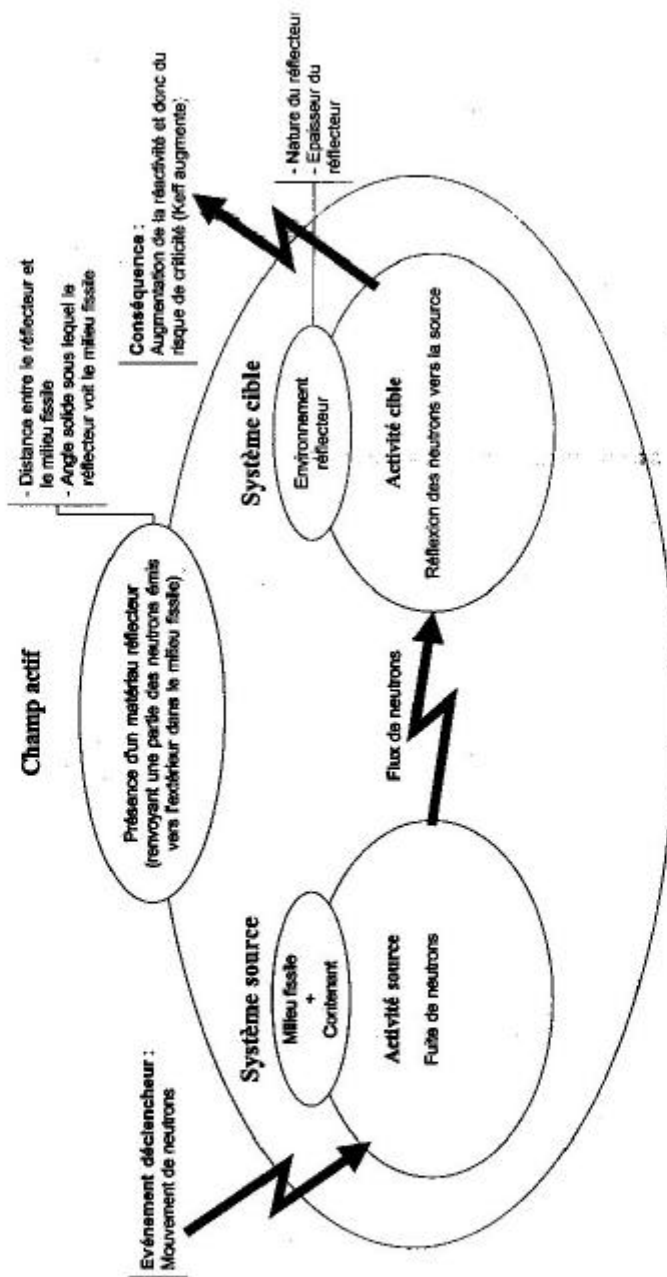
- Publication, pp. 11-35, 2001.
- [Quintana & al. 92] : Quintana Y., Kamel M., Lo A., "*Graph-Based Retrieval of Information in Hypertext systems*", in proceedings of the Tenth International Conference on Systems Documentation, ACM Press, pp. 157-168, 1992.
- [Ranwez 00] : notes personnelles prises lors de la soutenance de thèse de Sylvie Ranwez, Décembre 2000.
- [Roger & al. 98] : Roger D., Lavandier J., Kolmayer E., "*Navigation et interfaces : cartes conceptuelles et autres outils*", analyse bibliographique réalisée pour PARINFO (Programme d'Aide à la Recherche en INfOrmation), révision du 27 Avril 2000, disponible en ligne : <http://www.enssib.fr/bibliotheque/documents/travaux/kolmayer/kolm-navig0.html>.
- [Rouet & Tricot 95], Rouet J-F., Tricot A., "*Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive*", Sciences et techniques éducatives, vol. 2, n° 3, Hermès, Paris, 1995.
- [Rouet & Tricot 98] : Rouet J-F., Tricot A., "*Chercher de l'information dans un hypertexte: vers un modèle des processus cognitifs*", in Les hypermédias, approches cognitives et ergonomiques, Hermès, Paris, pp. 57-74, 1998.
- [Rouet 98] : Rouet J-F., Introduction aux PréActes du Quatrième Colloque Hypermédias et Apprentissages, Poitiers, 15-17 octobre 1998.
- [Roussey & al. 99] : Roussey C., Calabretto S., Pinon J-M., "*État de l'art en indexation et recherche d'information*", Document Numérique, Hermès, vol. 3, n° 3-4, pp. 121-149, Paris, 1999.
- [Sakarovitch 84] : Sakarovitch M., "*Programmation discrète. Optimisation combinatoire, Méthodes mathématiques et algorithmiques*", Hermann, Paris 1984.
- [Salminen & al. 95] : Salminen, A., Tague-Sutcliffe J., Mcclellan C., "*From Text to Hypertext by Indexing*", in ACM Transactions on Information Systems, vol. 13, n° 1, pp. 69-99, janvier 1995.
- [Salton & Mc Gill 83] : Salton G., Mc Gill M-J. "*Introduction to modern information retrieval*", McGraw-Hill, pp. 157-198, 1983.
- [Savoy 96] : Savoy J., "*Citation schemes in hypertext information retrieval*", in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 99-120, 1996.
- [Scapin & Pierret-Golbreich 90] : Scapin D-L., Pierret-Golbreich C., "*Towards a method for task description: MAD*" , in L. Berlinguer and D. Berthelette (Eds), Work in display units 89, Elsevier Science Publishers, North Holland, 1990.
- [Scapin 86] : Scapin D-L., "*Guide ergonomique de conception des interfaces homme-machine*", Rapport Technique INRIA n° 77, Rocquencourt, Octobre 1986.
- [Scapin 90] : Scapin D-L., Pierret-Golbreich C., "*Towards a method for task description: MAD*" , in L. Berlinguer and D. Berthelette (Eds), Work in display units 89, Elsevier Science Publishers, North Holland, 1990.
- [Scapin 97] : Scapin D-L., "*Ergonomie des logiciels Interactifs*", transparents issus du cours d'ergonomie du DEA SETI, CEA Saclay, 1997.
- [Schreiber & Wielinga 93] : Schreiber G., Wielinga B., "*Model construction*", in Schreiber

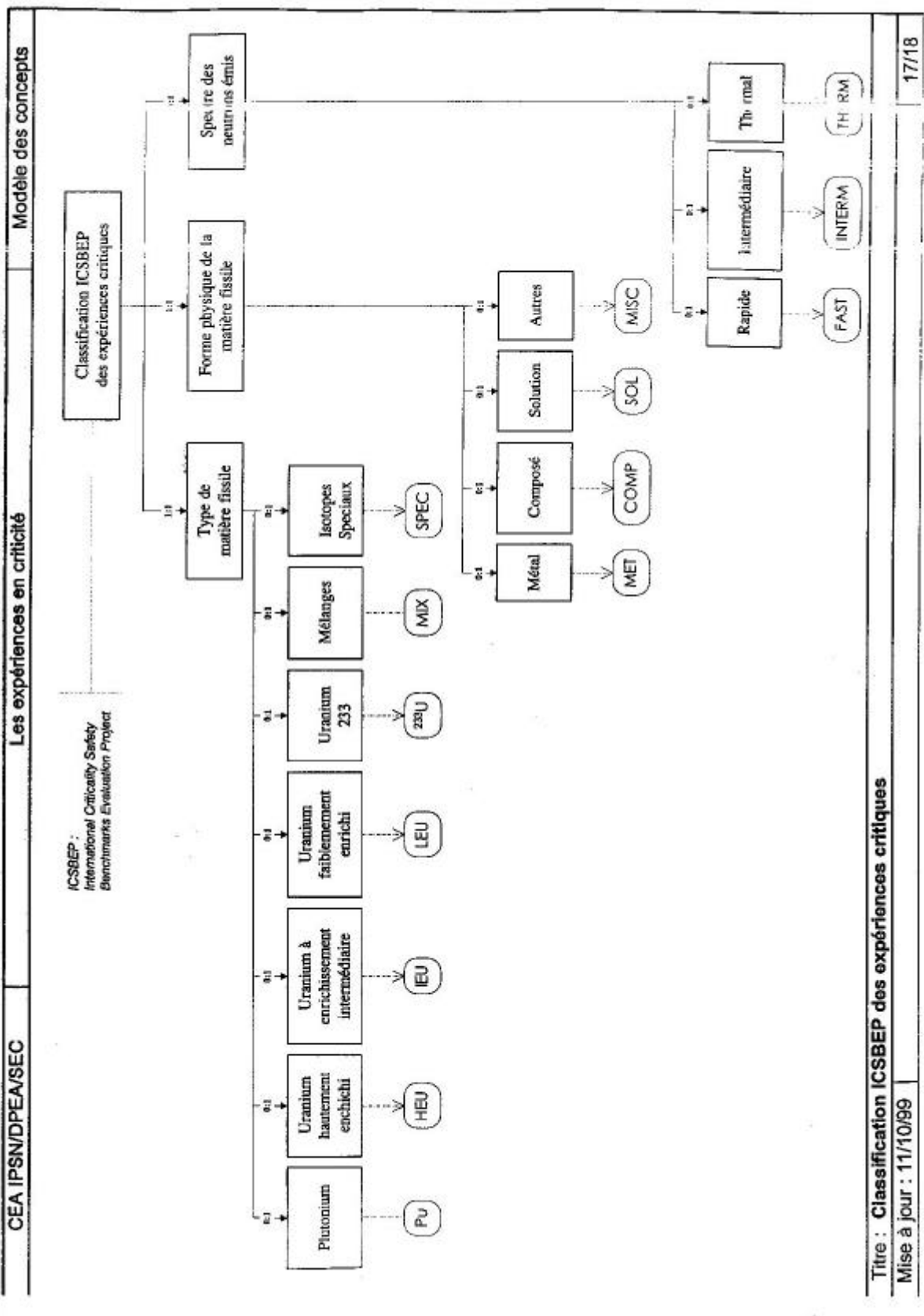
- G., Wielinga B., Breuker J. (Eds), *"Kads, a principled approach to knowledge-based systems development"*, Academic Press, 1993.
- [Sector 98] : notes personnelles prises lors de la visite de l'entreprise Sector, dans le cadre de la formation « Les Doctoriales », Juin 1998.
- [Shannon 48] : Shannon C. E. *"A mathematical Theory of Communication"*, The Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, juillet et octobre 1948.
- [Shapira & al. 96] : Shapira B., Shoval P., Raveh A., Hanani U., *"Hypertext browsing: a new model for information filtering based on user profiles and data clustering"*, Online & CDROM Review, vol. 20, n° 1, 1996.
- [Shepherd 89] : Shepherd A., *"Analysis and training in information technology tasks"*, in Task Analysis for Human-Computer Interaction, Dan Diaper ed., Ellis Horwood Ltd., pp. 15-55, GB, 1989.
- [Shneiderman 92] : Shneiderman B., *"Designing the User Interface: Strategies for Effective Human-Computer Interaction"*, 2^{ème} édition, Addison-Wesley, Reading, 1992.
- [Simoni & Fluhr 97] : Simoni J-L., Fluhr C., *"Accès à l'information à travers les graphes de termes"*, communication à ISKO-France, Lille, octobre 1997.
- [Simoni 00] : Simoni J-L., *"Accès à l'information à l'aide d'un graphe de termes construit automatiquement, intégration de l'interrogation et de la navigation"*, thèse de doctorat, Université Paris 7, Janvier 2000.
- [Smeaton 96] : Smeaton A. F., *"An overview of information retrieval"*, in Information Retrieval and Hypertext, Agosti M., Smeaton A. Eds., Kluwer Academic Publications, pp. 3-26, 1996.
- [Smith 96] : Smith W-J., *"ISO and ANSI Ergonomics Standards for Computer Products: A Guide to Implementation and Compliance"*, Prentice-Hall, 1996.
- [Tardieu & al. 83] : Tardieu H., Rochefeld A., Coletti C., *"La méthode Merise"*, Éditions d'organisation, 1983.
- [Taylor 62] : Taylor R., *"The process of asking questions"*, American documentation, pp. 391-397, 1962.
- [Tricot & al. 98] : Tricot A., Pierre-Demarcy C., El Boussarghini R., *"Définition d'aides en fonction des types d'apprentissages dans les environnements hypermédia"*, in Quatrième Colloque Hypermedias et Apprentissages, Poitiers, pp. 41-58, 15-17 octobre 1998.
- [Tricot & Nanard 98] : Tricot A., Nanard J., *"Un point sur la modélisation des tâches de recherche d'information dans le domaine des hypermédias"*, in Les hypermédias, approches cognitives et ergonomiques, A. Tricot & J-F. Rouet Eds., Hermès, Paris, pp. 35-56, 1998.
- [Van Campenhoudt 96] : Van Campenhoudt M., *"Abrégé de terminologie multilingue"*, centre de recherche Termisti, disponible en ligne : <http://www.refer.fr/termisti/theoweb3.htm>, 1996.
- [Van Craeynest & al. 00] : Van Craeynest J-M., Charlot J-M., Malavieille D., Bigeon P., Chaillot M., *"Traçabilité des connaissances avec MKSM"*, Actes du congrès scientifique NîmesTIC'2000, Nîmes, pp. 12-22, Septembre 2000.
- [Vignaux 94] : Vignaux G. : *"Les sciences cognitives : une introduction"*, Le Livre de Poche, La découverte, Paris, 1994.

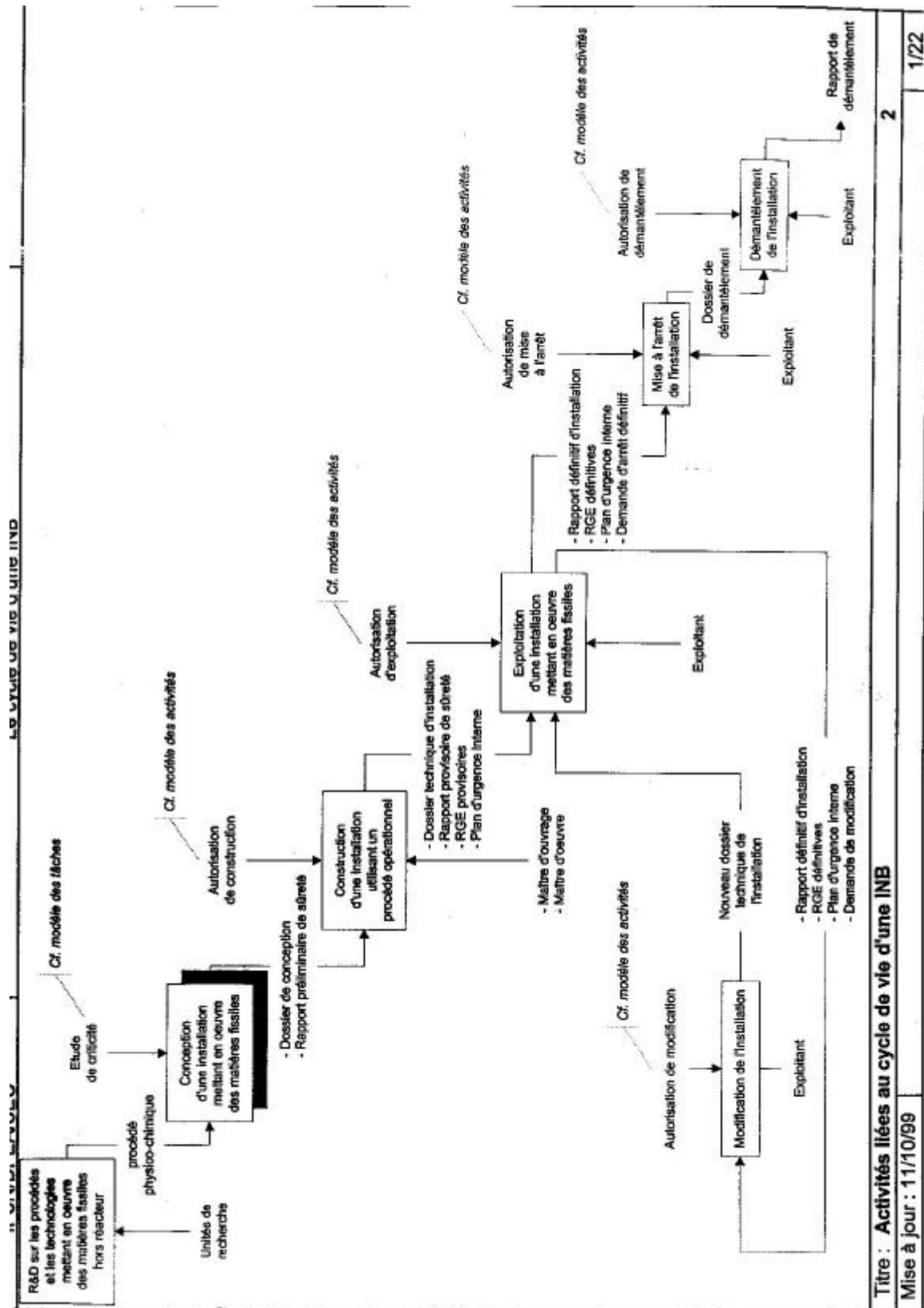
- [Vogel 89] : Vogel. C., *"Génie cognitif"*, Collection Sciences Cognitives, Masson, Paris, 1989.
- [Von Bertalanffy 68] : Von Bertalanffy, L., *"General Systems Theory"*, Foundation, Development, Applications, New York: G. Braziller (1968), Trad. française: *"Théorie générale des systèmes"*, Paris: Dunod, 1973.
- [Von Hippel 88] : Von Hippel, E., *"The Sources of Innovation"*, New York, NY: Oxford University Press, 1988.
- [Vora & Helander 97] : Vora P. R., Helander M. G., *"Hypertext and its Implications for the Internet"*, Handbook of Human-Computer Interaction, M. Helander, T. K. Landhauer, P. Prabhu eds., Elsevier Science B.V., pp.877-914, 1997.
- [Vora 98] : Vora P., *"Human factors methodology for designing web sites"*, in Human factors and web development, Forsythe C., Grose E. & Ratner J. (Eds.), Mahwah, New Jersey, Lawrence Erlbaum Associates, pp. 153-172, 1998.
- [Vossen & al. 89] : Vossen P., Meijs W., Broeder M., *"Meaning and structure in dictionary definitions"*, Boguraev B., Briscoe T., éditeurs, Computational lexicography for Natural language Processing, pp. 171-190. Longman Group UK limited, London, 1999.
- [Waern & Hägglund 97] : Wærn, Y., Hägglund S., *"User Aspects Of Knowledge-Based Systems"*, Handbook of Human-Computer Interaction, M. Helander, T.K. Landauer, P. Prabhu (eds.), Elsevier Science, 1997, pp. 1159-1176.
- [Wahlster & Kobsa 89] : Wahlster W., Kobsa A., *"User Models in Dialog Systems"*, A. Kobsa and W. Wahlster eds., Berlin, Springer-Verlag, pp. 4-34, 1989.
- [Wielinga & al. 93] : Wielinga B., Schreiber G., Breuker J., *"Modelling expertise"*, in Schreiber G., Wielinga B., Breuker J. (Eds.), *"Kads, a principled approach to knowledge-based systems development"*, Academic Press, 1993.
- [Woodhead 90] : Woodhead N., *"Hypertext & Hypermedia. Theory and Applications"*, Sigma Press, GB, 1991.
- [Wüster 76] : Wüster E., *"La théorie générale de la terminologie -un domaine interdisciplinaire impliquant la linguistique, la logique, l'ontologie, l'informatique et les sciences des objets "*, Actes du colloque international de terminologie (Québec, Manoir du lac Delage, 5-8 octobre 1975) , Essai de définition de la terminologie, H. DUPUIS éd., Québec, Régie de la langue française, p. 49-57, 1976.
- [Zarri & al. 99] : Zarri G. P., Bertino E., Black B., *"CONCERTO, An Environment for the "Intelligent" Indexing, Querying and Retrieval of Digital Documents"*, in proceedings of the 11th International Symposium ISMIS'99, Foundation of Intelligent Systems, LNAI, 1609, Varsovie, pp. 226-234, Juin 1999.
- [Zizi & Beaudoin-Lafon] : Zizi M., Beaudoin-Lafon M., *"Accessing Hyperdocuments through Interactive Dynamic Maps"*, in proceedings of the ACM ECHT'94, European Conference on Hypermedia Technologies, pp. 126-135, septembre 1994.

Annexe 2. Quelques diagrammes de connaissances issus du projet-pilote

Le projet opérationnel présenté dans cette thèse s'appuie sur la modélisation des connaissances effectuée en sûreté-criticité. Nous présentons ci-après quelques diagrammes issus de ce projet, dans le double but de permettre au lecteur de se représenter l'activité modélisée dans ce domaine et de se familiariser avec les formalismes utilisés. Les diagrammes des processus physiques et des concepts présentés se situent au niveau général du domaine de la maîtrise du risque de sûreté-criticité. Les diagrammes d'activités, de tâches et des objets sont spécifiques au sous-domaine des études de criticité pour la conception d'installations nucléaires de base hors réacteurs nucléaires.





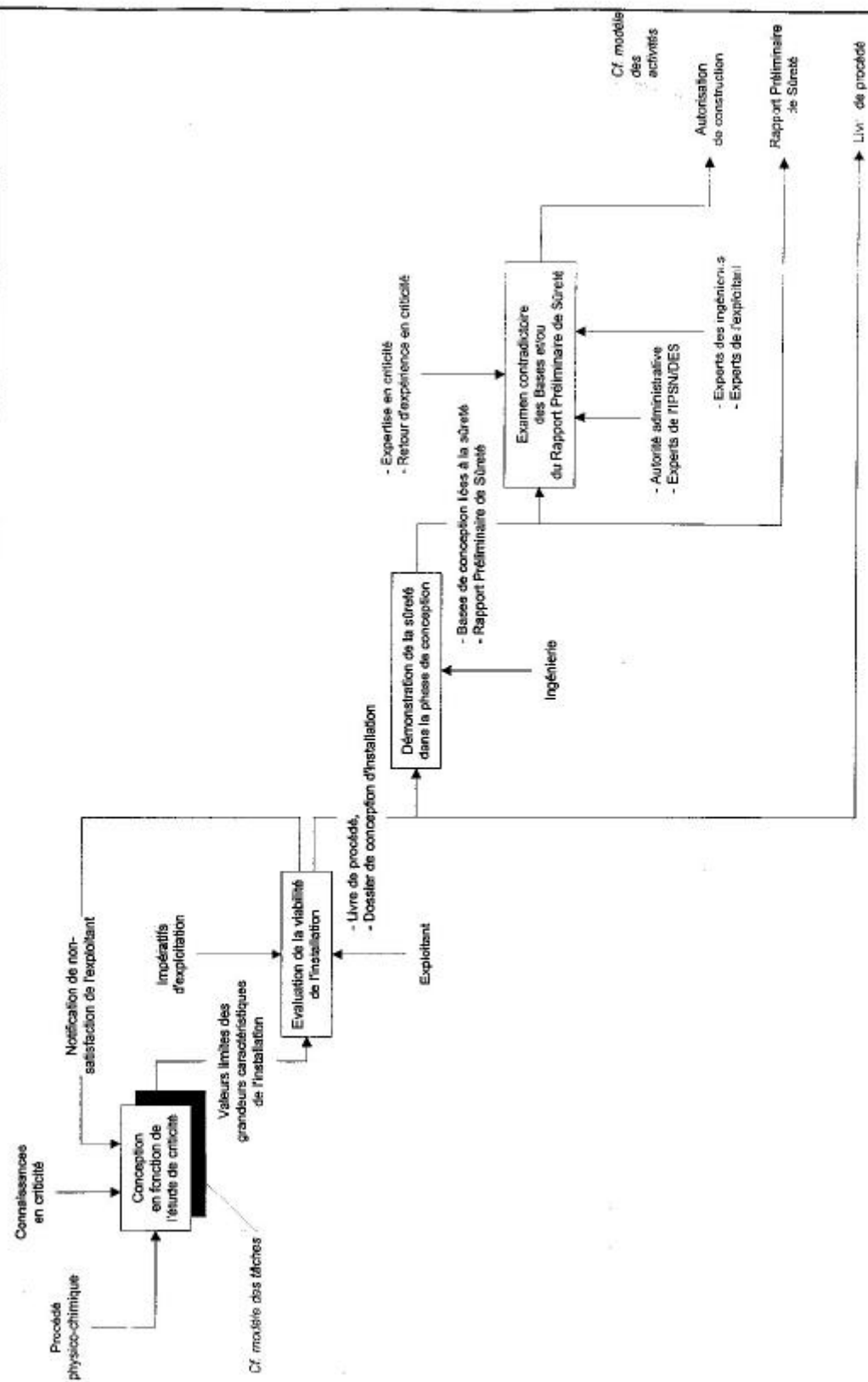


Titre : Activités liées au cycle de vie d'une INB

Mise à jour : 11/10/99

2

1/22

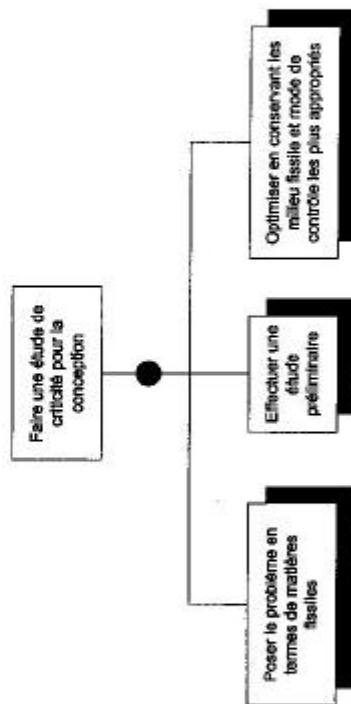


Titre : Conception d'une installation mettant en oeuvre des matières fissiles

Mise à jour : 11/10/99

3

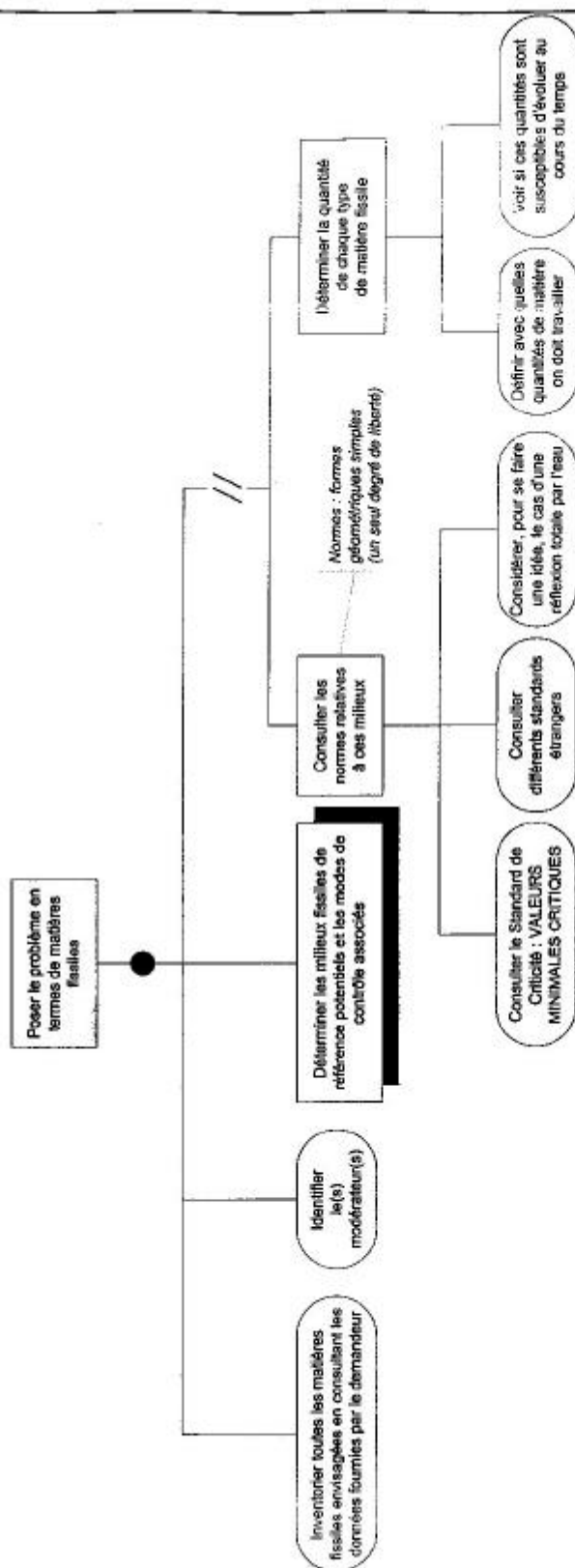
2/22



Titre : **Faire une étude de criticité pour la conception**

Mise à jour : 11/10/99

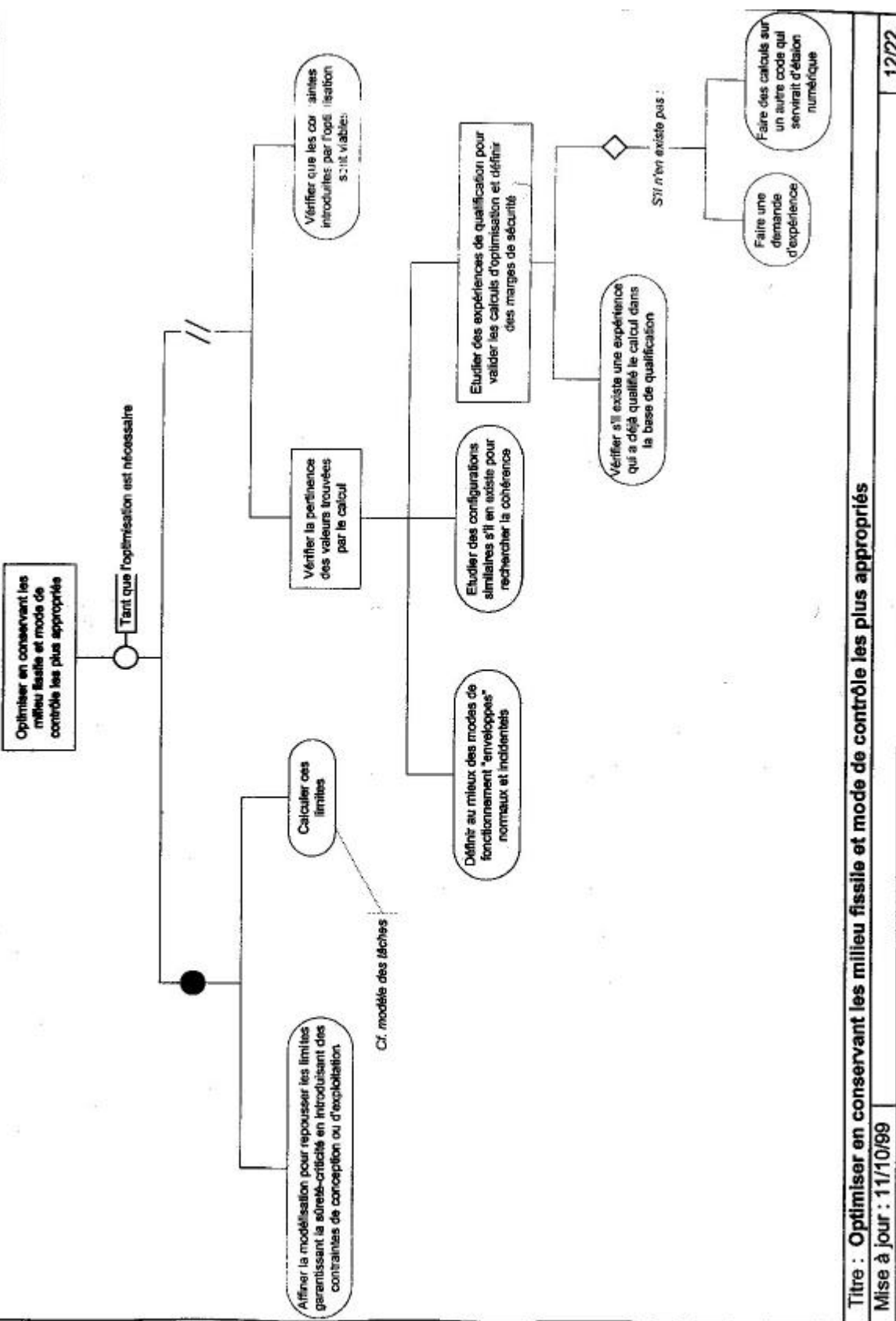
8/22

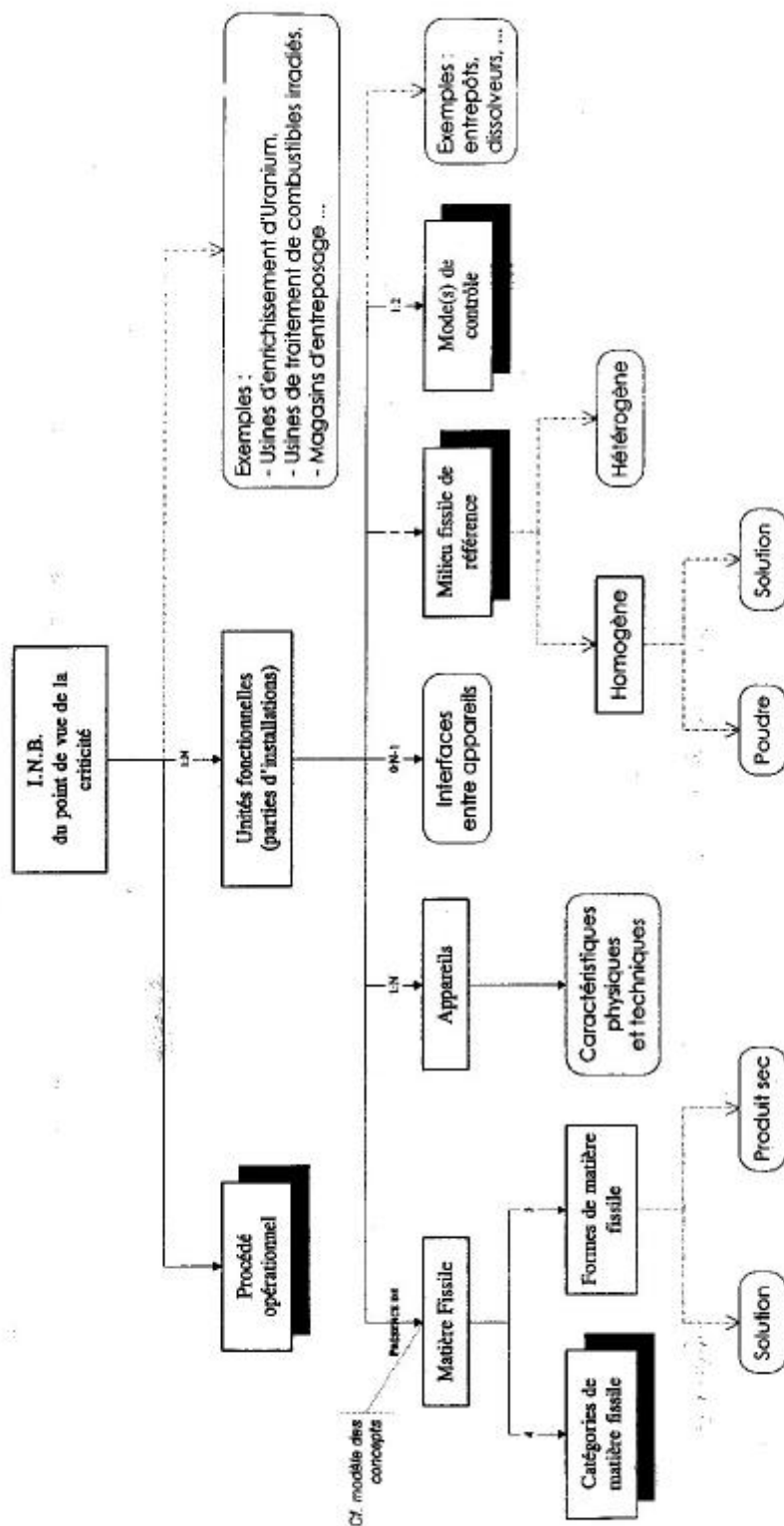


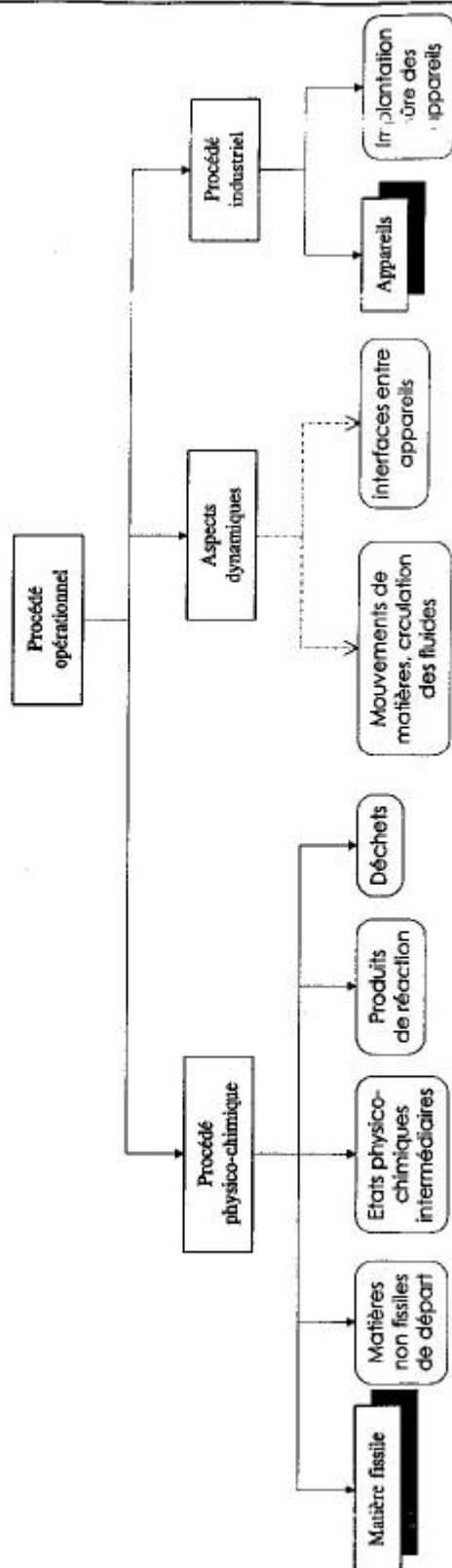
Titre : Poser le problème en termes de matières fissiles

Mise à jour : 11/10/99

9/22







Titre : Procédé opérationnel

Mise à jour : 11/10/99

16/22

Annexe 3. Fichiers XML sources du prototype

Cette annexe présente des fichiers XML complets réalisés pour un diagramme de connaissances et pour un profil d'utilisateur. Ces fichiers sont décrits en détail au chapitre 4.

Annexe 3.1. Fichier XML représentant un diagramme MKSM dans le LMCE

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE diagramme SYSTEM "diagramme.dtd">
<?xml-stylesheet type="text/xsl" href="Diagramme.xsl"?>
<DIAGRAMME>
  <HEAD>
    <TITLE>Activités liées au cycle de vie d'une INB</TITLE>
    <REDACTEUR>Lionel Médini</REDACTEUR>
    <PROPRIETAIRE>IPSN/DPEA/SEC</PROPRIETAIRE>
    <FORMALISME>Activite</FORMALISME>
    <SOUSDOMAINES>INB</SOUSDOMAINES>
  </HEAD>
  <IMG SRC="INB-Activite2.jpg" ALT="Page-2" BORDER="0 "
  USEMAP="#visImageMap"/>
  <MAP NAME="visImageMap">

    <NOEUD>
      <ID>activitedecomposee3</ID>
      <TYPE>activitedecomposee</TYPE>
      <TEXTE>Conception d'une installation mettant en oeuvre des matières
fissiles</TEXTE>
      <SHAPE>polygon</SHAPE>
      <COORDS>154,189,267,189,267,121,154,121,154,189</COORDS>
      <DIVSTYLE>left:270; top:90</DIVSTYLE>
      <LIEN>
        <TEXTE>Détailier l'activité</TEXTE>
        <DEST>INB-activite2.xml</DEST>
      </LIEN>
      <LIEN>
        <TEXTE>Liste des matières fissiles : modèle des concepts</TEXTE>
        <DEST>INB-concept4.xml</DEST>
      </LIEN>
      <RECHERCHE>
        <MIR>
          <NOM>Acrobat Search</NOM>
          <REFO>conception NEAR (installation OR INB) AND "matiere
fissile"</REFO>
        </MIR>
        <MIR>
          <NOM>SPIRIT</NOM>
          <REFO/>
        </MIR>
      </RECHERCHE>
    </NOEUD>

    <NOEUD>
      <ID>acteur35</ID>
      <TYPE>acteur</TYPE>
      <TEXTE>Unités de recherche</TEXTE>
      <SHAPE>polygon</SHAPE>
      <COORDS>33,177,93,177,93,139,33,139,33,177</COORDS>
      <DIVSTYLE>left:94; top:120</DIVSTYLE>
      <RECHERCHE>
        <MIR>
          <NOM>Acrobat Search</NOM>
          <REFO>"unite de recherche" OR "unites de recherche" OR
laboratoire</REFO>
        </MIR>
      </RECHERCHE>
    </NOEUD>
  </MAP>
</DIAGRAMME>
```

```

    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>activite27</ID>
  <TYPE>activite</TYPE>
  <TEXTE>R et D sur les procédés et les technologies mettant en oeuvre des
matières fissiles hors réacteur</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>3,79,124,79,124,3,3,3,3,79</COORDS>
  <DIVSTYLE>left:125; top:0</DIVSTYLE>
  <LIEN>
    <TEXTE>Liste des matières fissiles : modèle des concepts</TEXTE>
    <DEST>INB-concept4.xml</DEST>
  </LIEN>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>recherche AND developpement AND procede AND technologie AND
"matiere fissile"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot43</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Etude de criticité</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>180,73,241,73,241,35,180,35,180,73</COORDS>
  <DIVSTYLE>left:242; top:0</DIVSTYLE>
  <LIEN>
    <TEXTE>Faire une étude de criticité : modèle des tâches</TEXTE>
    <DEST>INB-tachel.xml</DEST>
  </LIEN>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"etude de criticite"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>activite47</ID>
  <TYPE>activite</TYPE>
  <TEXTE>Construction d'une installation utilisant un procédé
opérationnel</TEXTE>
  <SHAPE>polygon</SHAPE>

```

```

<COORDS>313,312,426,312,426,243,313,243,313,312</COORDS>
<DIVSTYLE>left:425; top:190</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>(Construction NEAR installation) AND procede</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>acteur</ID>
  <TYPE>acteur</TYPE>
  <TEXTE>Exploitant</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>869,679,929,679,929,648,869,648,869,679</COORDS>
  <DIVSTYLE>left:930; top:600</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>exploitant</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot1</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Procédé physico-chimique</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>86,119,184,119,184,81,86,81,86,119</COORDS>
  <DIVSTYLE>left:184; top:10</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"procede physico-chimique"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot9a</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Rapport définitif d'installation</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>313,631,472,631,472,614,313,614,313,631</COORDS>
  <DIVSTYLE>left:470; top:570</DIVSTYLE>
  <RECHERCHE>
    <MIR>

```

```

    <NOM>Acrobat Search</NOM>
    <REFO>"rapport définitif d installation"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot9b</ID>
  <TYPE>flot</TYPE>
  <TEXTE>RGE définitives</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>313,644,472,644,472,632,313,632,313,644</COORDS>
  <DIVSTYLE>left:470; top:570</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"RGE définitives"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot9c</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Plan d'urgence interne</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>313,657,472,657,472,645,313,645,313,657</COORDS>
  <DIVSTYLE>left:470; top:570</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"plan d urgence interne"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot9d</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Demande de modification</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>313,682,472,682,472,658,313,658,313,682</COORDS>
  <DIVSTYLE>left:470; top:570</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"demande de modification"</REFO>
    </MIR>
    <MIR>

```

```

    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>activite9</ID>
  <TYPE>activite</TYPE>
  <TEXTE>Modification de l'installation</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>128,595,271,595,271,565,128,565,128,595</COORDS>
  <DIVSTYLE>left:271; top:505</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>modification NEAR installation</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot56</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Nouveau dossier technique de l'installation</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>313,607,411,607,411,554,313,554,313,607</COORDS>
  <DIVSTYLE>left:411; top:525</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"dossier technique" NEAR (installation OR INB)</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>acteur46</ID>
  <TYPE>acteur</TYPE>
  <TEXTE>Exploitant</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>169,643,229,643,229,613,169,613,169,643</COORDS>
  <DIVSTYLE>left:229; top:595</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>exploitant</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>

```



```

</NOEUD>

<NOEUD>
  <ID>flot20</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Autorisation de modification</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>127,539,271,539,271,509,127,509,127,539</COORDS>
  <DIVSTYLE>left:271; top:450</DIVSTYLE>
  <LIEN>
    <TEXTE>Procédure générique d'attribution des autorisations : modèle des
    activités</TEXTE>
    <DEST>INB-activite8.xml</DEST>
  </LIEN>
</NOEUD>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"autorisation de modification"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot49</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Exploitant</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>753,631,814,631,814,601,753,601,753,631</COORDS>
  <DIVSTYLE>left:814; top:550</DIVSTYLE>
</NOEUD>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"exploitant"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot48</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Rapport de démantèlement</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>978,675,1069,675,1069,637,978,637,978,675</COORDS>
  <DIVSTYLE>left:945; top:540</DIVSTYLE>
</NOEUD>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"rapport de demantelement"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>

```

```

    </RECHERCHE>
  </NOEUD>

  <NOEUD>
    <ID>flot27</ID>
    <TYPE>flot</TYPE>
    <TEXTE>Autorisation de démantèlement</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>853,508,944,508,944,471,853,471,853,508</COORDS>
    <DIVSTYLE>left:950; top:470</DIVSTYLE>
    <LIEN>
      <TEXTE>Procédure générique d'attribution des autorisations : modèle des
    activités</TEXTE>
      <DEST>INB-activite8.xml</DEST>
    </LIEN>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"autorisation de demantelement"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

  <NOEUD>
    <ID>flot25</ID>
    <TYPE>flot</TYPE>
    <TEXTE>Autorisation de mise à l'arrêt</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>749,431,817,431,817,378,749,378,749,431 </COORDS>
    <DIVSTYLE>left:817; top:341</DIVSTYLE>
    <LIEN>
      <TEXTE>Procédure générique d'attribution des autorisations : modèle des
    activités</TEXTE>
      <DEST>INB-activite8.xml</DEST>
    </LIEN>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"autorisation de mise a l arret"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

  <NOEUD>
    <ID>flot17</ID>
    <TYPE>flot</TYPE>
    <TEXTE>Dossier de démantèlement</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>789,576,880,576,880,539,789,539,789,576</COORDS>
    <DIVSTYLE>left:880; top:490</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>

```

```

    <REFO>"dossier de demanletelement"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot16</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Autorisation d'exploitation</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>509,312,585,312,585,274,509,274,509,312</COORDS>
  <DIVSTYLE>left:585; top:240</DIVSTYLE>
  <LIEN>
    <TEXTE>Procédure générique d'attribution des autorisations : modèle des
    activités</TEXTE>
    <DEST>INB-activite8.xml</DEST>
  </LIEN>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"autorisation d exploitation"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Autorisation de construction</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>324,183,415,183,415,145,324,145,324,183</COORDS>
  <DIVSTYLE>left:415; top:105</DIVSTYLE>
  <LIEN>
    <TEXTE>Procédure générique d'attribution des autorisations : modèle des
    activités</TEXTE>
    <DEST>INB-activite8.xml</DEST>
  </LIEN>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"autorisation de construction"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot18a</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Rapport définitif d'installation</TEXTE>

```

```

<SHAPE>polygon</SHAPE>
<COORDS>621,447,780,447,780,431,621,431,621,447</COORDS>
<DIVSTYLE>left:775; top:430</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"rapport definitif d installation"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot18b</ID>
  <TYPE>flot</TYPE>
  <TEXTE>RGE définitives</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>621,460,780,460,780,448,621,448,621,460</COORDS>
  <DIVSTYLE>left:775; top:430</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"RGE definitives"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot18c</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Plan d'urgence interne</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>621,473,780,473,780,461,621,461,621,473</COORDS>
  <DIVSTYLE>left:775; top:430</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"plan d urgence interne"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot18d</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Demande d'arrêt définitif</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>621,499,780,499,780,474,621,474,621,499</COORDS>
  <DIVSTYLE>left:775; top:430</DIVSTYLE>
  <RECHERCHE>

```

```

<MIR>
  <NOM>Acrobat Search</NOM>
  <REFO>"demande d arret definitif"</REFO>
</MIR>
<MIR>
  <NOM>SPIRIT</NOM>
  <REFO/>
</MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot57a</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Dossier technique d'installation</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>373,332,540,332,540,315,373,315,373,332</COORDS>
  <DIVSTYLE>left:545; top:280</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"dossier technique d installation"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot57b</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Rapport provisoire de sûreté</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>373,345,540,345,540,333,373,333,373,345</COORDS>
  <DIVSTYLE>left:545; top:280</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"rapport provisoire de surete"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot57c</ID>
  <TYPE>flot</TYPE>
  <TEXTE>RGE provisoires</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>373,358,540,358,540,346,373,346,373,358</COORDS>
  <DIVSTYLE>left:545; top:280</DIVSTYLE>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"RGE provisoires"</REFO>
  </MIR>

```

```

    <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot57d</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Plan d'urgence interne</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>373,383,540,383,540,359,373,359,373,383</COORDS>
  <DIVSTYLE>left:545; top:280</DIVSTYLE>
  <RECHERCHE>
    <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"plan d urgence interne"</REFO>
  </MIR>
  <MIR>
  <NOM>SPIRIT</NOM>
  <REFO/>
  </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot53a</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Dossier de conception</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>199,216,366,216,366,202,199,202,199,216</COORDS>
  <DIVSTYLE>left:367; top:140</DIVSTYLE>
  <RECHERCHE>
    <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"dossier de conception"</REFO>
  </MIR>
  <MIR>
  <NOM>SPIRIT</NOM>
  <REFO/>
  </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>flot53b</ID>
  <TYPE>flot</TYPE>
  <TEXTE>Rapport preliminaire de sûreté</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>199,240,366,240,366,217,199,217,199,240</COORDS>
  <DIVSTYLE>left:367; top:140</DIVSTYLE>
  <RECHERCHE>
    <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"rapport preliminaire de surete"</REFO>
  </MIR>
  <MIR>
  <NOM>SPIRIT</NOM>
  <REFO/>
  </MIR>

```

```

    </RECHERCHE>
  </NOEUD>

  <NOEUD>
    <ID>activite34</ID>
    <TYPE>actvite</TYPE>
    <TEXTE>Démantèlement de l'installation</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>853,614,944,614,944,576,853,576,853,614</COORDS>
    <DIVSTYLE>left:944; top:530</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>demantelement NEAR installation</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

  <NOEUD>
    <ID>activite33</ID>
    <TYPE>actvite</TYPE>
    <TEXTE>Mise à l'arrêt de l'installation</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>742,537,825,537,825,499,742,499,742,537</COORDS>
    <DIVSTYLE>left:825; top:450</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"Mise a l arret de l installation"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

  <NOEUD>
    <ID>acteur15</ID>
    <TYPE>acteur</TYPE>
    <TEXTE>Exploitant</TEXTE>
    <SHAPE>polygon</SHAPE>
    <COORDS>517,561,577,561,577,531,517,531,517,561</COORDS>
    <DIVSTYLE>left:577; top:480</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>Exploitant</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

  <NOEUD>

```

```

<ID>activite4</ID>
<TYPE>actvite</TYPE>
<TEXTE>Exploitation d'une installation mettant en oeuvre des matières
fissiles</TEXTE>
<SHAPE>polygon</SHAPE>
<COORDS>490,455,604,455,604,387,490,387,490,455</COORDS>
<DIVSTYLE>left:604; top:325</DIVSTYLE>
<LIEN>
  <TEXTE>Liste des matières fissiles : modèle des concepts</TEXTE>
  <DEST>INB-concept4.xml</DEST>
</LIEN>
<RECHERCHE>
  <MIR>
    <NOM>Acrobat Search</NOM>
    <REFO>"Exploitation NEAR (installation OR INB) AND "matieres
fissiles"</REFO>
  </MIR>
  <MIR>
    <NOM>SPIRIT</NOM>
    <REFO/>
  </MIR>
</RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>acteur11a</ID>
  <TYPE>acteur</TYPE>
  <TEXTE>Maître d'ouvrage</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>320,424,419,424,419,410,320,410,320,424</COORDS>
  <DIVSTYLE>left:419; top:360</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"maitre d ouvrage"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

<NOEUD>
  <ID>acteur11b</ID>
  <TYPE>acteur</TYPE>
  <TEXTE>Maître d'oeuvre</TEXTE>
  <SHAPE>polygon</SHAPE>
  <COORDS>320,448,419,448,419,425,320,425,320,448</COORDS>
  <DIVSTYLE>left:419; top:360</DIVSTYLE>
  <RECHERCHE>
    <MIR>
      <NOM>Acrobat Search</NOM>
      <REFO>"maitre d oeuvre"</REFO>
    </MIR>
    <MIR>
      <NOM>SPIRIT</NOM>
      <REFO/>
    </MIR>
  </RECHERCHE>
</NOEUD>

```


</MAP>
</DIAGRAMME>

Annexe 3.2. Fichier XML représentant un profil d'utilisateur dans le LMCE

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE profil SYSTEM "Profil.dtd">
<PROFIL>

  <UTILISATEUR>
    <NOM>Médini</NOM>
    <PRENOM>Lionel</PRENOM>
    <PASSWORD>toto</PASSWORD>
  </UTILISATEUR>

  <EXPERTISE>
    <INB>&expert;</INB>
    <TRANSPORTS>&expert;</TRANSPORTS>
    <EXPERIENCES>&apprenant;</EXPERIENCES>
    <CODES>&apprenant;</CODES>
    <STANDARDS>&novice;</STANDARDS>
    <DIFFUSION>&novice;</DIFFUSION>
  </EXPERTISE>

  <INFORMATION>
    <STRATEGIE>&cumulative;</STRATEGIE>
    <FILTRAGE>&large;</FILTRAGE>
  </INFORMATION>

  <VISUALISATION>
    <CADRES>
      <DIAGRAMME>
        <SHOW>12</SHOW>
        <WIDTH>80%</WIDTH>
        <HEIGHT>80%</HEIGHT>
        <XPOS>20%</XPOS>
        <YPOS>10%</YPOS>
      </DIAGRAMME>

      <TITREDIAGRAMME>
        <SHOW>&OUI;</SHOW>
        <WIDTH>80%</WIDTH>
        <HEIGHT>10%</HEIGHT>
        <XPOS>20%</XPOS>
        <YPOS>0%</YPOS>
      </TITREDIAGRAMME>

      <REPNESRI>
        <SHOW>&PEUTETRE;</SHOW>
        <WIDTH>20%</WIDTH>
        <HEIGHT>50%</HEIGHT>
        <XPOS>0%</XPOS>
        <YPOS>0%</YPOS>
      </REPNESRI>

      <DOCUMENT>
        <SHOW>&PEUTETRE;</SHOW>
        <WIDTH>20%</WIDTH>
        <HEIGHT>50%</HEIGHT>
        <XPOS>0%</XPOS>
        <YPOS>0%</YPOS>
      </DOCUMENT>
    </CADRES>
  </VISUALISATION>
</PROFIL>
```

```

<ARBREFONCTIONNEL>
  <SHOW>&NON ; </SHOW>
  <WIDTH>20%</WIDTH>
  <HEIGHT>50%</HEIGHT>
  <XPOS>0%</XPOS>
  <YPOS>50%</YPOS>
</ARBREFONCTIONNEL>

<DICOCONCEPT>
  <SHOW>&OUI ; </SHOW>
  <WIDTH>20%</WIDTH>
  <HEIGHT>50%</HEIGHT>
  <XPOS>0%</XPOS>
  <YPOS>50%</YPOS>
</DICOCONCEPT>

<DICOACTEUR>
  <SHOW>&NON ; </SHOW>
  <WIDTH>20%</WIDTH>
  <HEIGHT>50%</HEIGHT>
  <XPOS>0%</XPOS>
  <YPOS>50%</YPOS>
</DICOACTEUR>

<CADDIECONCEPTS>
  <SHOW>&OUI ; </SHOW>
  <WIDTH>20%</WIDTH>
  <HEIGHT>50%</HEIGHT>
  <XPOS>0%</XPOS>
  <YPOS>0%</YPOS>
</CADDIECONCEPTS>

<INFOPROFIL>
  <SHOW>&NON ; </SHOW>
  <WIDTH/>
  <HEIGHT/>
  <XPOS/>
  <YPOS/>
</INFOPROFIL>
</CADRES>
</VISUALISATION>

</PROFIL>

```