



HAL
open science

Système de recherche d'information sémantique pour les bases de visioconférences médicales à travers les graphes conceptuels

Ameni Yengui

► To cite this version:

Ameni Yengui. Système de recherche d'information sémantique pour les bases de visioconférences médicales à travers les graphes conceptuels. Recherche d'information [cs.IR]. Faculté des sciences économiques et de gestion Sfax, 2016. Français. NNT: . tel-01767195

HAL Id: tel-01767195

<https://theses.hal.science/tel-01767195v1>

Submitted on 15 Apr 2018

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

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

République Tunisienne
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Sfax
Faculté des Sciences Economiques et de Gestion de Sfax
MIRACL - Sfax



Thèse pour l'obtention du titre de docteur en

Informatique

Systeme de recherche d'information sémantique pour les bases de visioconférences médicales à travers les graphes conceptuels

Présentée et soutenue publiquement par

Ameni YENGUI

Le 15 Juin 2016

Membres du jury

Mme. Lamia BELGUITH	Professeur, FSEGS-Sfax	Président
Mr. Mahmoud NEJI	Maître de conférences, FSEGS-Sfax	Directeur de thèse
Mr. Mohamed Adel ALIMI	Professeur, ENIS-Sfax	Rapporteur
Mr. Mounir BEN AYED	Maître de conférences, FSS-Sfax	Rapporteur
Mme. Ikram AMOUS	Maître de conférences, ENET'COM-Sfax	Membre

Année Universitaire 2015 - 2016

***Système de recherche d'information
sémantique pour les bases de
visioconférences médicales à travers les
graphes conceptuels***

« On se lasse de tout, sauf de comprendre. »

Virgile.

Résumé

Notre travail s'intègre dans le cadre de la recherche d'information (RI) d'un type particulier de documents audiovisuels, à savoir « les visioconférences ».

Dans cette thèse, nous présentons notre système de recherche d'information OSSVIRI. Il est composé de trois modules. Dans le module d'analyse, nous avons proposé une technique statistique exploitant les fréquences des mots afin d'extraire les termes simples, composés et spécifiques des documents. Dans le module d'indexation, nous avons utilisé le modèle ontologique afin d'associer les termes à leurs concepts, d'extraire les relations entre eux à partir des documents, de désambigüiser les concepts reconnus afin de mieux couvrir la sémantique du contenu documentaire. Les concepts et les relations sont représentés sous forme d'un graphe conceptuel. Dans le module de recherche, nous proposons d'un côté, une technique de reformulation des requêtes utilisateurs basée sur les ressources externes et les profils des utilisateurs et de l'autre côté une technique d'appariement basée sur l'expansion combinée de requêtes et des documents guidées par le contexte du besoin en information et des contenus documentaires.

Ce système est validé en utilisant les métriques issues du domaine de la recherche d'information et des comparaisons avec une approche statistique existante. Nous montrons ainsi que notre approche obtient de bons résultats.

MOTS-CLÉS : recherche d'information, indexation sémantique, analyse, reformulation requête, appariement requête-document, ontologie, thésaurus, visioconférence, mesure statistique

Abstract

Our work fits into the context of information retrieval (IR) of a particular type of audiovisual documents, namely "video conferencing."

In this thesis, we introduce our OSSVIRI information retrieval system, which composed of three modules. In the analysis module, we have proposed a statistical technique exploiting the word frequency in order to extract the simple, compound and specific terms from the documents.

In the indexing module, we used the ontology to associate the terms with their concepts, retrieve the relations between them and disambiguate the concepts to improve the semantic content of the documents. The concepts and relations are represented as a conceptual graph. In the research module, we have proposed a technique of users' query reformulation based on external resources and users' profiles and a technique of pairing based on the combined expansion of the requests and the documents guided by the context of the requirement in information and the documentary contents.

This system is validated using the metrics from the research information and comparisons with existing statistical approach. We show that our approach achieves good results.

KEYWORDS: information retrieval, semantic indexing, analysis, query reformulation, pairing query-document, ontology, thesaurus, videoconference, statistical measure.

Je dédie cette thèse

à mes parents

à ma petite famille

à ma sœur

à mon frère et sa femme

à toute la famille Yengui

à la famille Lahmédi

à Nadia

*à tous ceux qui m'ont toujours
encouragé*

Remerciements

Je tiens avant tout à remercier Mr. Mahmoud NEJI pour son soutien, son écoute et ses commentaires précieux tout au long de mes travaux de thèse. Son avis et sa direction ont largement contribué à la bonne rédaction de mon manuscrit de thèse. Merci pour les nombreux échanges durant mes phases de rédaction et merci pour l'ensemble de discussions agréables que nous avons pu avoir.

Je remercie Mr. Rafik BOUAZIZ pour m'avoir proposé de travailler sur ce sujet de recherche passionnant. Je le suis hautement reconnaissante d'avoir pu partager son expérience et ses qualités de recherche. J'espère avoir tiré profit pour mes futures années de chercheur. Ses connaissances à la fois théoriques et pratiques, associées à son dynamisme scientifique m'ont impressionnée tout au long de mes années d'étude. Merci d'être toujours disponible pour me conseiller.

Merci à Mr. Mohamed Adel ALIMI, Professeur à l'Université de Sfax, ainsi qu'à Mr. Mounir BEN AYED, Maître de Conférences à l'Université de Sfax, d'avoir accepté d'être rapporteurs de cette thèse, et pour l'honneur qu'ils me font en participant au jury. Merci également à Mme Lamia BELGUITH, Professeur à l'Université de Sfax et à Mme Ikram AMOUS, Maître de Conférences à l'Université de Sfax, d'avoir accepté de juger ce travail et de faire partie du jury. Je tiens à leur exprimer toute ma reconnaissance pour l'intérêt porté à ce travail.

Mes remerciements vont aussi à tous ceux qui m'ont permis d'effectuer ce travail dans des bonnes conditions. Merci à tous mes amis d'avoir été là et particulièrement ma chère Nadia.

Merci à mes parents pour m'avoir encouragé et aidé à mener ma vie d'étudiant si long temps. Votre soutien constant à travers ces longues années m'a été précieux.

Je remercie enfin ma petite famille : merci Adnen et Iyed d'être là. Merci pour m'avoir accompagné et encouragé durant des moments pas toujours facile.

Table des matières

Introduction générale.....	1
I.Contexte et problématique	1
II. Objectif de la thèse.....	3
III. Plan de la thèse.....	3

Partie 1 État de l’art : Modélisation, Indexation et Recherche Audiovisuelle

Chapitre 1 : Indexation et recherche dans les Documents Audiovisuels	8
Introduction	9
I.Catégories d’indexation	10
I.1. L’indexation manuelle	10
I.2. L’indexation semi-automatique	10
I.3. L’indexation automatique	11
I.4. L’opposition entre l’indexation manuelle, semi-automatique et automatique.....	11
I.5. Objectifs	11
II. Types d’indexation.....	12
II.1. Indexation par concept.....	12
II.2. Indexation par contenu	14
II.2.1. Indexation de texte.....	15
II.2.2.Indexation des images fixes.....	17
II.2.3.Indexation de l’audio	20
II.2.4.Indexation de la vidéo.....	21
II.3. Indexation par le contexte.....	22
II.4. Indexation par le contenu et le contexte	23
III. Fossé sémantique et fusion multimodale	24
IV. Les descripteurs.....	27
IV.1. Les descripteurs de niveau bas.....	27
IV.1.1. Descripteurs image de bas niveau	27
IV.1.2. Descripteurs audio de bas niveau	27
IV.1.3. Descripteurs vidéo de bas niveau	28
IV.2. Les descripteurs de niveau moyen	28
IV.2.1. Descripteurs image de niveau sémantique moyen	29
IV.2.2. Descripteurs audio de niveau sémantique moyen	29
IV.2.3. Descripteurs vidéo de niveau sémantique moyen	29
IV.3. Les descripteurs de niveau haut	29
IV.3.1. Descripteurs image de niveau sémantique haut	30

IV.3.2. Descripteurs audio de niveau sémantique haut	30
IV.3.3. Descripteurs vidéo de niveau sémantique haut	30
V. Module de recherche d'information	30
Conclusion	31
Chapitre 2 : Formalisme de représentation des documents et base de connaissances	33
Introduction	34
I. Les graphes conceptuels (GCs).....	34
I.1. Généralités	34
I.2. Notations	35
I.3. Quelques définitions	35
I.4. Les graphes simples	37
I.5. Les graphes emboîtés typés	38
I.6. Les règles et contraintes	39
I.7. Synthèse	40
II. Les ontologies	40
II.1. Définitions	40
II.1.1. Vocabulaire contrôlé.....	40
II.1.2. Taxonomie	41
II.1.3. Thésaurus.....	41
II.1.4. Réseau sémantique	41
II.2. De la philosophie à l'informatique	42
II.2.1. Intelligence artificielle.....	42
II.2.2. Web sémantique	42
II.3. Les niveaux des ontologies.....	43
II.4. Composants des ontologies	44
II.5. Langages de spécification d'ontologie.....	44
II.5.1. Resource Description Framework : RDF	44
II.5.2. DAML+OIL.....	45
II.5.3. Ontology Web Language (OWL)	45
II.6. Ontologies et Recherche d'Information	45
Conclusion	48
Chapitre 3 : Recherche d'information	49
Introduction	50
I. Reformulation des requêtes.....	50
I.1. Approches de reformulation	53
I.1.1. Reformulation automatique de requête.....	53
I.1.2. Reformulation interactive de requête.....	55
I.2. Utilité du profil utilisateur	55
I.2.1 Les techniques de construction des profils utilisateurs.....	56

I.2.2. Les techniques d'exploitation des profils utilisateurs	57
I.3. Expansion automatique des requêtes	57
I.4. Combinaison des présentations des requêtes	58
I.5. Réinjection de pertinence.....	58
II. Appariement document-requête	59
II.1. Modèles d'appariement	59
II.1.1. Le modèle booléen.....	59
II.1.2. Le modèle booléen pondéré.....	60
II.1.3. Le modèle vectoriel	60
II.1.4. Le modèle vectoriel étendu.....	61
II.1.5. Le modèle probabiliste	63
II.1.6. Modèle inférentiel.....	64
II.1.7. Le modèle différentiel bayésien	64
II.1.8. Les modèles de langage	65
II.2. Projection sémantique.....	66
II.3. Calcul de correspondance	66
II.4. Classement.....	67
II.5. Présentation	67
Conclusion	68
Discussion.....	69

Partie 2 Contribution: OSSVIRI - Ontology-based System for Semantic Video-conference Information Retrieval and Indexation

Chapitre 1 : Modélisation structurelle et sémantique.....	73
I. Structure générale de notre SRI	74
II. Schéma de modélisation du contenu visioconférence.....	75
III. Modélisation structurelle : multi-facette	78
III.1. La notion de facette (ou vue).....	78
III.2. Modélisation du contenu visuel.....	78
III.2.1. La facette perceptive	78
III.2.2. La facette structurelle	79
III.2.3. La facette spatiale.....	79
III.2.4. La facette symbolique	80
III.2.5. Facette événementielle	82
III.2.6. Facette temporelle	83
III.2.7. Facette signal.....	85
III.3. Modélisation du contenu audio	94
III.3.1. Spécification conceptuelle et description du contenu audio	94
III.3.2. Segmentation audio et structure du document	95

III.3.3 Modélisation de la sous-facette audio	95
IV. Modélisation sémantique	97
IV.1. Modélisation du contenu visuel : Facette sémantique	97
IV.1.1. Modèle de la sous- facette visuelle	97
IV.1.2. Extraction de concepts visuels	97
IV.2. Modélisation du contenu Audio : extraction des termes	97
IV.2.1. Transcription de la parole	98
IV.2.2. Extraction des termes simples.....	102
IV.2.3. Extraction des Termes composés.....	106
Conclusion	110
Chapitre 2 : Indexation du contenu des visioconférences médicales.....	112
Introduction	113
I. Techniques utilisées	114
I.1. Ontologie dense et thésaurus	114
I.2. Choix d'ontologie de domaine.....	114
I.2.1. OntoMénélas	115
I.2.2. UMLS	115
I.3. Formalisme de représentation	116
II. Module d'indexation.....	118
II.1. Extraction des Concepts	118
II.1.1. Choix de la ressource externe	122
II.1.2. Ambigüités des termes.....	123
II.1.3. Pondération des concepts.....	123
II.1.4. Algorithme d'extraction des concepts	124
II.2. Extraction des relations sémantiques entre les concepts	125
II.3. Mise à jour de la base d'index	126
III. Représentation sous forme d'un graphe unique	130
V. Réseaux des graphes	133
Conclusion	133
Chapitre 3 : Recherche d'information dans une base de visioconférences médicales ...	135
Introduction	136
I.Processus de reformulation	136
I.1. Phase de capture du contexte statique.....	138
I.2. Phase de reformulation	138
II. Étape d'appariement document-requête.....	142
II.1. Appariement de graphes	143
II.1.1. Projection de graphes.....	143
II.1.2. Calcul de similarité.....	144
II.2. Calcul de pertinence	146

Conclusion	147
------------------	-----

Partie 3 Réalisation : Expérimentation et évaluation

Chapitre 1 : Implémentation..... 150

Introduction	151
I. Récapitulatif des apports méthodologiques de notre thèse	151
II. Expérimentation.....	152
II.1. Environnement de développement	152
II.1.1.Choix du Système d’exploitation	152
II.1.2. Choix de l’outil d’exploitation de l’ontologie	152
II.1.3. Choix du Langage de programmation	152
II.1.5. Choix du SGBD.....	153
II.2. Description du système OSSVIRI	153
II.2.1. Menu Fichier.....	154
II.2.2. Module d’Analyse	155
II.2.2. Module d’Indexation	157
Conclusion	162

Chapitre 2 : Création du corpus et évaluation 163

Introduction	164
I.Descriptif du protocole d’expérimentation	164
II. Construction du corpus	165
II.1. Corpus du test	165
II.2. Corpus d’appui	165
II.3. Ressources externes	166
II.3.1.Présentation	166
II.3.2. Les concepts dans les ressources	166
II.3.3. Les relations entre les concepts et les types sémantiques.....	166
II.3.4. Les relations sémantiques entre les types sémantiques	166
III. Expérimentations et évaluations	167
III.1. Système OSSVIRI.....	168
III.2. Méthodologie d’évaluation	168
III.2.1. Description générale de la méthode d’évaluation	168
III.2.2. Mesure d’évaluation.....	168
III.3. Module « Analyse »	170
III.3.1. Extraction des termes simples par élimination des mots vides	170
III.3.2. Extraction des termes composés : détermination du seuil IMA.....	171
III.4. Module « Indexation »	172
III.4.1. Extraction des concepts.....	172
III.4.2. Extraction des relations sémantiques	173

III.4.3. Traitement des ambiguïtés	174
III.5. Module « Recherche ».....	175
III.5.1. Reformulation des requêtes.....	175
III.5.2. Appariement Document-Requête.....	176
Conclusion	181
Conclusion et perspectives.....	183
I.Contribution.....	183
I.1. Sur le plan théorique	183
I.1.1. Module d'analyse.....	183
I.1.2. Module d'indexation.....	184
I.1.3. Module de recherche.....	184
I.2. Sur le plan pratique et technique.....	185
II. Perspectives.....	186
II.1. Perspectives à court terme	186
II.2. Perspectives à moyen terme	186
II.3. Perspectives à long terme	187
Nos Publications	188
Bibliographie.....	190
Glossaire.....	206
Annexes	209

Liste des Figures

Figure 1. 1: Comparaison entre l'indexation classique et l'indexation sémantique.....	16
Figure 1. 2 : Facettes d'une image selon le modèle EMIR	19
Figure 1. 3 : Les étapes dans le module de recherche d'information.....	31
Figure 1. 4: Exemple de représentation avec les GCs.....	35
Figure 1. 5 : (a) Exemple de treillis de concepts, (b) exemple de treillis de relations	35
Figure 1. 6: Exemple d'opération de projection de GCs.....	37
Figure 1. 7 : SG description du profil du médecin Baccar Hédi	37
Figure 1. 8: Un graphe emboîté typé.....	39
Figure 1. 9: Exemple de règle de SG	40
Figure 1. 10: Les types d'ontologies	44
Figure 1. 11 : Principe d'intégration d'une ontologie au module de RI, [Said L'hadj, 09]	46
Figure 1. 12 : Utilisation d'ontologie pour le module d'indexation.....	47
Figure 1. 13 : Module de recherche.....	50
Figure 1. 14 : Modèle d'augmentation, [Fuhr <i>et al.</i> , 03].....	63
Figure 1. 15 : Projection sémantique.....	66
Figure 1. 16: Présentation.....	68
Figure 2. 1: Architecture générale de notre système	75
Figure 2. 2: Structure d'une visioconférence	76
Figure 2. 3: Description événementielle	83
Figure 2. 4 : Les relations temporelles d'Allen	85
Figure 2. 5: Classification de relations conceptuelles	90
Figure 2. 6: Les mouvements de caméra.....	92
Figure 2. 7: Description générique du contenu audio	95
Figure 2. 8: Treillis des concepts audio.....	96
Figure 2. 9 : Treillis de concepts visuels.....	98
Figure 2. 10 : Module de modélisation (Analyse).....	99
Figure 2. 11: Etapes de l'indexation audio.....	100
Figure 2. 12: description générique du contenu audio	102
Figure 2. 13 : Module d'indexation.....	119
Figure 2. 14 : Exemple d'un concept décrit par DOE.....	120
Figure 2. 15 : Exemple de concept d'une ontologie.....	121

Figure 2. 16: Mise à jour d'Index.....	130
Figure 2. 17: Exemple de représentation formalisme de GC	131
Figure 2. 18 : Description générique du contenu audio	131
Figure 2. 19: Reformulation de la requête utilisateur.....	137
Figure 2. 20 : Phase de récupération du contexte statique des utilisateurs	138
Figure 2. 21 : Architecture du processus de reformulation	140
Figure 2. 22: Opération de projection des GCs.....	143
Figure 3. 1 : Interface Module Analyse d'OSSVIRI.....	153
Figure 3. 2 : Interface Module Indexation d'OSSVIRI.....	154
Figure 3. 3 : Interface Module Recherche d'OSSVIRI.....	154
Figure 3. 4 : Menu Fichier.....	154
Figure 3. 5 : Extraction des termes.....	155
Figure 3. 6 : Extraction des termes spécifiques.....	155
Figure 3. 7 : Liste des termes simples	156
Figure 3. 8 : Liste des termes composés.....	156
Figure 3. 9 : Liste des termes spécifiques	157
Figure 3. 10 : Liste des termes	157
Figure 3. 11 : Menu « Indexation ».....	158
Figure 3. 12 : Ontologie affiché sous la forme arborescente	158
Figure 3. 13 : Liste des relations sémantique entre les concepts.....	159
Figure 3. 14 : Liste des concepts	159
Figure 3. 15 : Index sous forme de graphe conceptuel.....	159
Figure 3. 16 : Fenêtre de suppression d'un document dans la base d'index.....	160
Figure 3. 17 : Identification de l'utilisateur.....	160
Figure 3. 18 : Connexion de l'utilisateur.....	161
Figure 3. 20 : Menu recherche	161
Figure 3. 19 : Inscription de l'utilisateur.....	161
Figure 3. 21 : Résultat du module de recherche.....	162
Figure 3. 22 : Le protocole d'expérimentation.....	164
Figure 3. 23 : Exemple d'un document XML	165
Figure 3. 24 : Architecture de OSSVIRI.....	169
Figure 3. 25 : Variation de MAP en fonction du seuil de l'IMA	171
Figure 3. 26 : Variation de P@5 en fonction du seuil de l'IMA.....	172
Figure 3. 27 : Courbe de la précision à 11 points de rappel.....	172

Figure 3. 28 : Courbe de la précision à 11 points de rappel : concepts / concepts + relations	174
Figure 3. 29 : Courbe de la précision à 11 points de rappel : sans et avec traitements des ambiguïtés	175
Figure 3. 30 : Courbe précision-rappel en 11 points pour Cosinus et Sim	179
Figure 3. 31 : Comparaison Cosinus/Sim des moyennes des précisions n par requête.....	179
Figure 3. 32 : seuil relatif en fonction du nombre de termes de la requête	180

Liste des Tableaux

Tableau 1. 1: exemples de projets de recherche dans l'indexation multimédia et la création de moteur de recherche, [Benmokhtar, 10].....	25
Tableau 2. 1 : Description des différents types de mouvement	93
Tableau 2.2 : Concepts associés	122
Tableau 2.3 : Concepts dénotés.....	122
Tableau 2. 4: Exemples de modification de requête	137
Tableau 2. 5: Extraction des concepts	142
Tableau 3. 1: Exemple de concept C0001365	166
Tableau 3. 2: Le concept C1516400 et ses types sémantiques.....	167
Tableau 3. 3: Une relation sémantique entre deux types sémantiques.....	167
Tableau 3. 4: Valeur de la MAP et de la P@5 en fonction du seuil de l'IMA.....	171
Tableau 3. 5: Résultat en MAP et de la P@5 pour les deux systèmes	173
Tableau 3. 6: Résultat en MAP et de la P@5 pour les deux scénarios : concepts / concepts + relations	174
Tableau 3. 7: Résultat en MAP et de la P@5 sans et avec traitements des ambiguïtés	174
Tableau 3. 8: Quelques résultats de la désambiguïsation.....	175
Tableau 3. 9: Résultat en Rappel et Précision sans et avec reformulation des requêtes	176
Tableau 3. 10: Précision à 5, à 10, à 30 documents restitués, R-précision moyenne et F à 5, à 10, à 30 documents pour <i>Sim</i> et <i>Cosinus</i>	178
Tableau 3. 11: Evaluation de la pertinence par jugement de l'utilisateur	181
Tableau 3. 12: Comparaison de l'efficacité des deux types de classement.....	181

Liste des Algorithmes

Algorithme 1 : Extraction des Termes Simples	106
Algorithme 2 : Extraction des Termes composés.....	110
Algorithme 3 : Extraction des concepts	125
Algorithme 4 : Extraction des relations sémantiques entre concepts.....	126
Algorithme 5 : Prise en compte de l'ajout d'un nouveau document.....	127
Algorithme 6 : Prise en compte de la suppression d'un nouveau document.....	128
Algorithme 7 : Modification d'un document : ajout d'information	129
Algorithme 8: Modification d'un document : suppression d'information	129
Algorithme 9 : Reformulation d'une requête	140
Algorithme 10 : Extraction des concepts	141
Algorithme 11 : Appariement Document-Requête	146
Algorithme 12 : Pertinence Document-Requête	147

Liste des Extraits

Extrait 1: Exemple de concept dans OntoMénélas.....	116
Extrait 2: Extrait XML du concept C1516400	120
Extrait 3: Extrait XML du concept C1516400	121
Extrait 4 : Extrait XML d'une visioconférence médicale	125

Introduction générale

I. Contexte et problématique

Avec le développement de l'Internet, les technologies de l'information et de communication (TIC) se sont développées. Elles constituent la somme de trois secteurs : le secteur informatique, le secteur électronique et le secteur de télécommunication. Ce dernier regroupe les technologies de communication comme par exemple les didacticiels, les espaces numériques d'apprentissage, les espaces numériques de travail et les visioconférences qui constituent une technologie multimédia de communication. La multiplication du nombre des données visioconférences a engendré un fort besoin dans les domaines d'éducation, d'archivage ou de distribution de contenu visioconférence. Cette diversité a abouti à une croissance des besoins des utilisateurs en information. En effet l'utilisateur ne se contente plus de subir l'information en spectateur ; il navigue, il cherche, il trouve, il compare et il échange les informations qu'il rencontre. L'accès aux informations stockées dans une base documentaire avec une gigantesque masse de données devient alors une tâche pénible. Afin de traiter ce problème, une discipline toute entière est née. Cette discipline est appelée **R**echerche d'**I**nformation (RI). Elle s'intéresse au développement des techniques et des outils qui permettent de retrouver une information intéressante qui satisfait le besoin de l'utilisateur en informations pertinentes. Ces outils sont appelés des **S**ystèmes de **R**echerche d'**I**nformation (SRI). Ainsi, un SRI permet de sélectionner parmi un volume d'information, les informations pertinentes qui répondent au besoin en information. Dans ce système, le besoin en information est exprimé sous forme de requête.

Dans un SRI, chaque document est montré par une représentation intermédiaire. Cette dernière est directement exploitée par les SRI. Elle décrit le contenu du document par des unités significatives dans le document appelées des descripteurs. Cette description est appelée l'indexation du document. De la même manière, une requête est décrite par un ensemble de descripteurs. Pour retrouver les documents vis-à-vis d'une requête, le SRI compare la représentation de cette requête avec la représentation de chaque document. Cette comparaison est réalisée au moyen d'une fonction de correspondance et un score de pertinence est affecté à chaque document. Ces scores permettent de représenter à l'utilisateur les documents les plus pertinents. Le SRI est donc composé de trois modules : un module d'analyse, un module d'indexation et un module de recherche. Ce dernier comporte deux tâches : la reformulation de requête et l'appariement.

Partant de l'idée que pour un utilisateur de SRI, ce qui compte le plus c'est la précision et la pertinence des réponses fournies par sa requête, nous cherchons à mettre en place une analyse

conceptuelle du contenu audiovisuel. Cette analyse devra permettre d'intégrer les différents types de média (image, texte et audio) et aussi de spécifier la manière dont le document audiovisuel sera indexé. Elle a pour but de permettre une description fine et précise du contenu.

À l'heure actuelle, plusieurs propositions ont été faites pour l'indexation et la recherche par le contenu des documents vidéo mais nous remarquons qu'aucune proposition n'a traité les documents visioconférences. Ces documents se caractérisent par leur structure spécifique différente des autres types de documents audiovisuels.

Dans le cadre de nos travaux, nous nous intéressons aux aspects modélisation, indexation et recherche du contenu des documents visioconférences.

Un autre problème important pour l'efficacité des modules d'indexation et de recherche, et en raison en particulier de la nature visuelle du signal visioconférence, les données visioconférences sont perçues et interprétées différemment par des personnes différentes. La problématique de la représentation du document visioconférence, de son contenu sémantique et structurel se retrouve principalement lors des phases d'indexation, d'expression des requêtes et d'interaction avec l'utilisateur.

L'indexation basée sur les ontologies comporte un certain nombre d'avantages par rapport aux méthodes traditionnelles d'indexation (seulement basées sur des mots clés). En effet, Les termes qui sont responsables de la sélection d'un document ne sont pas forcément ceux de la requête (ils peuvent être des synonymes ou des termes décrivant des concepts plus spécifiques). Notre travail s'inscrit dans l'optique d'utilisation des ontologies dans le domaine de la médecine. Ces ontologies seront exploitées dans un système d'indexation de documents audiovisuels.

Le module de recherche se base essentiellement sur la requête exprimée par l'utilisateur pour répondre à ses besoins. En effet, quel que soit le système de recherche utilisé, le résultat de recherche ne peut être pertinent si la requête ne décrit pas explicitement et clairement les besoins de l'utilisateur. Or, il est généralement reconnu que l'utilisateur se contente de donner quelques mots clés. Ces derniers sont issus d'une connaissance générale sur le sujet recherché. Par conséquent, les documents renvoyés par le système peuvent ne pas satisfaire les besoins de l'utilisateur. La reformulation des requêtes est l'une des stratégies qui permet d'améliorer ses constructions. Elle consiste de manière générale à enrichir la requête utilisateur en ajoutant des termes permettant de mieux exprimer le besoin.

Les travaux décrits dans cette thèse s'intéressent à la création d'un SRI pour les documents visioconférences tout en traitant l'aspect sémantique de ces documents et en utilisant des ontologies de domaine médical. Les visioconférences sont utilisées dans différents domaines comme le domaine éducatif, médical, commercial. Et vue que le domaine médical est un domaine actif et peut être utilisé pour l'enseignement et vue l'existence de plusieurs ontologies qui traitent ce domaine, nous l'avons choisi comme domaine d'application dans nos travaux de recherche.

II. Objectif de la thèse

Ce travail s'inscrit dans un contexte d'indexation et de recherche d'information multimédia. Dans la perspective d'améliorer l'efficacité des systèmes de recherche d'information visioconférence en termes de rappel et de précision, nous proposons une modélisation du contenu sémantique de la visioconférence. Cette modélisation permet d'intégrer les descriptions issues de différents types de média et l'aspect temporel de la visioconférence. Les interprétations du contenu d'un document visioconférence sont souvent sémantiquement très riches et variées. Ceci rend l'indexation plus compliquée lorsqu'il s'agit d'une simple indexation soit par des mots-clés du fait qu'on doit choisir les meilleurs index pour décrire un contenu très riche en information, soit en exploitant des descripteurs de bas niveau qui ne reflètent pas les descripteurs sémantiques du contenu visioconférence. La même difficulté est aussi trouvée lors de module de recherche. En effet, l'utilisation de « sac de mots-clés » pour rechercher un segment vidéo ne permet pas d'avoir des réponses pertinentes. Pour surmonter ces difficultés, nous proposons de modéliser le contenu d'un document visioconférence en exploitant la notion de concept et de relation conceptuelle. Cette proposition permettra de faciliter l'indexation et la recherche des documents visioconférences par le contenu sémantique. Ensuite, nous proposons d'indexer ces documents. Cette indexation permet de décrire le contenu de ces documents par des descripteurs sémantiques. Ainsi, notre travail consiste à extraire ces descripteurs à partir de ces documents. Il s'agit de proposer une méthode d'extraction des concepts et des relations sémantiques entre ces concepts. Dans un premier temps, nous extrayons les termes simples et les termes composés à partir de la visioconférence médicale. Ces termes sont transformés par la suite en concepts. Dans cette étape de transformation, nous utilisons des ressources sémantiques externes. Ensuite, cette ressource est utilisée pour extraire les relations sémantiques entre les concepts. Le dernier processus consiste à reformuler la requête utilisateur et à calculer le score de correspondance entre la requête et les visioconférences médicales déjà indexées pour mieux satisfaire le besoin de l'utilisateur.

Notre objectif à travers ce travail est de développer un SRI basé sur des ontologies de domaine médical. Pour atteindre cet objectif, nous allons suivre les étapes suivantes :

- analyser le contenu du document audiovisuel pour le rendre accessible et exploitable par le module d'indexation ;
- exploiter les ontologies pour la création d'index formant une base d'index ;
- reformuler la requête utilisateur en utilisant des ressources externes ;
- calculer le score de correspondance entre requête / visioconférences.

III. Plan de la thèse

La suite de ce rapport est organisée en trois parties. La première partie contenant un état de l'art de notre domaine d'étude qui se compose de trois chapitres. La seconde partie décrit à la fois le choix des ontologies à utiliser dans notre système d'indexation et la phase de conception de notre SRI et la dernière partie présente le SRI développé, sa réalisation en

expliquant le choix des outils utilisés, sa description par des captures d'écran de ses différentes interfaces et son évaluation à l'aide en le comparant avec d'autres existants.

- **Partie 1- État de l'art : Modélisation, Indexation et Recherche Audiovisuelle**

- **Chapitre 1 : Indexation et recherche dans les Documents Audiovisuels**

Dans ce chapitre, nous présentons un panorama des travaux liés aux problèmes de l'accès aux documents audiovisuels et leur manipulation. Plus précisément, nous nous intéressons aux travaux sur l'indexation et la recherche par le contenu. L'objectif est de faire une idée sur quelques systèmes et approches proposés dans la littérature. Nous introduisons aussi à la fin de ce chapitre quelques outils et schémas de description audiovisuels

- **Chapitre 2 : Formalisme de représentation et base de connaissances**

L'objectif de ce chapitre est d'introduire le formalisme des graphes conceptuels et les ontologies. Ces dernières sont utilisées pour présenter des descriptions partagées et plus ou moins formelles d'un domaine permettant ainsi d'ajouter une couche sémantique aux systèmes informatiques. L'indexation des documents par des mots ou groupes de mots posent plusieurs problèmes.

Le formalisme de graphes conceptuels sera utilisé comme modèle opérationnel dans le cadre de notre proposition. Nous présentons aussi les différents termes liés aux ontologies et les différents niveaux des ontologies.

- **Chapitre 3 : Recherche d'information**

Dans ce chapitre, nous présentons des travaux traitant les problèmes de recherche d'information. Parmi ces travaux, nous trouvons ceux qui traitent la reformulation des requêtes. D'autres présentent les méthodes de calcul de correspondance et de similarité entre la requête reformulée et les documents.

- **Partie 2- Contribution : OSSVIRI - Ontology-based System for Semantic Visio-conference Information Retrieval and Indexation**

- **Chapitre 1 : Modélisation Structurelle et sémantique**

La modélisation du contenu de la visioconférence a pour objectif de formuler des descriptions basées sur ses caractéristiques audiovisuelles. Ces dernières sont généralement dépendantes du type média. En effet, pour le contenu audio par exemple, la description est basée sur l'information qu'il contient (parole, musique, ...). Pour fournir plusieurs niveaux d'abstraction en exploitation le contenu de la visioconférence, deux niveaux de modélisation (sémantique et structure) seront obligatoires

Ce chapitre détaille la modélisation structurelle et sémantique de notre proposition. Nous présentons l'ensemble des points liés au modèle proposé. Ce

schéma comporte une description multi-facette et multimodale du contenu audiovisuel.

- **Chapitre 2 : Indexation du contenu des visioconférences médicales**

La modélisation des connaissances est un sujet majeur pour différentes problématiques. Différentes méthodes, formalismes et outils ont été proposés et mis en œuvre pour modéliser, stocker et utiliser les informations au sein des bases de données ou des bases de connaissances pour répondre aux besoins d'un utilisateur jouant par la suite un rôle central dans l'interaction et le dialogue avec l'utilisateur tels que les graphes conceptuels et les ontologies.

Pour fournir plusieurs niveaux d'abstraction en exploitation le contenu d'une visioconférence, nous proposons un schéma de modélisation à deux niveaux : niveau structurel et niveau sémantique. Les descriptions du premier niveau sont calculées automatiquement en utilisant des descripteurs visuels. La modélisation sémantique est une abstraction qui permet de lier des descriptions de bas niveau avec le monde réel.

- **Chapitre 3 : Recherche d'information dans une base de visioconférences médicales**

Dans ce chapitre, nous détaillons notre proposition d'amélioration des performances des systèmes de recherche. Pour mieux satisfaire les besoins d'un utilisateur, deux processus seront nécessaires : la reformulation de la requête initiale de l'utilisateur et le calcul de correspondance entre l'index de la requête et la base d'index des documents visioconférences. Durant l'étape de reformulation, nous décrivons les trois étapes sous-jacentes : l'extraction et la sélection des termes, la pondération des termes et la réécriture de la requête. Puis nous calculons le score de correspondance entre la requête reformulée et les documents visioconférences médicales pour donner à l'utilisateur le document le plus pertinent.

- **Partie 3- Réalisation : expérimentation et évaluation**

La dernière partie de cette thèse est consacrée à l'expérimentation de notre SRI ainsi que l'évaluation de ces résultats. Afin de le tester, nous avons implémenté un outil nommé **O**ntology-based **S**ystem for **S**emantic **V**isio-conference **I**nformation **R**etrieval and **I**ndexation (OSSVIRI) qui regroupe les différentes phases de notre SRI.

- **Chapitre 1 : Implémentation**

Ce chapitre décrit l'implémentation de l'OSSVIRI qui présente un SRI pour les documents visioconférences médicales. Ce système est composé de trois modules : l'Analyse, l'Indexation et la Recherche.

- **Chapitre 2 : Création du corpus et évaluation**

Ce chapitre présente les expérimentations que nous avons réalisées. Ces expérimentations ont pour objectif d'évaluer notre approche d'analyse, d'indexation (extraction des concepts et des relations entre concepts) et de recherche (reformulation de requête et appariement). Dans ces expérimentations, nous appliquons notre approche à une base de documents visioconférences Medical Visio-conference Documents (MVD).

- **Conclusion générale**

Dans ce dernier chapitre, nous effectuons un résumé sur l'apport de cette thèse. Nous tirons aussi le bilan des réalisations et nous mentionnons les perspectives de recherche suggérées par ce travail.

Partie 1

*État de l'art : Modélisation,
Indexation et Recherche
Audiovisuelle*

Chapitre 1 : Indexation et recherche dans les Documents Audiovisuels

Dans ce chapitre, nous présentons un panorama des travaux liés aux problèmes de l'accès aux documents audiovisuels et de leur manipulation. Plus précisément, nous nous intéressons aux travaux traitant l'indexation et la recherche par le contenu. L'objectif est de faire une idée sur quelques systèmes et approches proposés dans la littérature. Nous introduisons aussi à la fin de ce chapitre quelques outils et schémas de description audiovisuels.

Introduction.....	8
I.Catégories d'indexation	10
I.1. L'indexation manuelle.....	10
I.2. L'indexation semi-automatique	10
I.3. L'indexation automatique.....	11
I.4. L'opposition entre l'indexation manuelle, semi-automatique et automatique	11
I.5. Objectifs.....	11
II.Types d'indexation	12
II.1. Indexation par concept	12
II.2. Indexation par contenu	14
II.2.1. Indexation de textes.....	15
II.2.2. Indexation d'images fixes.....	17
II.2.3. Indexation de l'audio.....	20
II.2.4. Indexation de la vidéo	21
II.3. Indexation par le contexte	22
II.4. Indexation par le contenu et le contexte	23
III. Fossé sémantique et fusion multimodal	24
IV. Les descripteurs	27
IV.1. Les descripteurs de niveau bas.....	27
IV.1.1. Descripteurs image de bas niveau.....	27
IV.1.2. Descripteurs audio de bas niveau.....	27
IV.1.3. Descripteurs vidéo de bas niveau.....	28
IV.2. Les descripteurs de niveau moyen	28
IV.2.1. Descripteurs image de niveau sémantique moyen	29
IV.2.2. Descripteurs audio de niveau sémantique moyen	29
IV.2.3. Descripteurs vidéo de niveau sémantique moyen	29
IV.3. Les descripteurs de niveau haut	29
IV.3.1. Descripteurs image de niveau sémantique haut	30
IV.3.2. Descripteurs audio de niveau sémantique haut	30
IV.3.3. Descripteurs vidéo de niveau sémantique haut	30
V. Module de recherche d'information	30
Conclusion.....	31

Introduction

De nombreuses ressources pédagogiques sont accessibles via des moteurs de recherche sur Internet mais celles-ci correspondent souvent à de simples présentations en ligne de documents qui n'ont pas été créés spécifiquement pour leur exploitation dans des environnements d'apprentissage.

Dans un contexte réel, lorsqu'un utilisateur souhaite rechercher des documents ou des portions de documents, il est souvent plus pratique pour lui d'utiliser une information sémantique (un événement, un concept spécifique) pour obtenir les réponses les plus pertinentes. Or les systèmes actuels ne satisfont pas vraiment ce besoin du fait que dans la plupart des cas, leurs auteurs privilégient un type spécifique de média. Par exemple, ils indexent la vidéo en exploitant uniquement leur aspect visuel. Il est difficile de traiter efficacement l'information sémantique par une telle approche.

L'indexation par le contenu permet d'assurer une gestion automatisée de grands corpus de documents multimédia et de fournir aux utilisateurs des moyens efficaces d'accès à ces documents. Les systèmes de recherche par le contenu indexent les média selon des descripteurs extraits automatiquement depuis le contenu des média et permettent de rechercher les média au sein d'une base de données en fonction de leurs caractéristiques visuelles. Mais, il est souvent difficile, en partant d'une analyse bas niveau du contenu média, d'atteindre le niveau sémantique. Partant de cette idée, et considérant aussi que pour un utilisateur d'un système de recherche d'information, ce qui compte le plus, c'est la précision et la pertinence des réponses fournies par sa requête. Plusieurs travaux d'indexation ont visé un comblement du fossé sémantique qui sépare les éléments de bas niveau (signal) et les éléments de haut niveau (informations et connaissances). Les caractéristiques sémantiques (ou de haut niveau) extraites automatiquement dans les documents multimédia s'appuient sur des algorithmes de classification automatique. Il s'agit alors d'inférer des connaissances sémantiques en apprenant dans la plupart des cas des régularités dans les caractéristiques de bas niveau. Beaucoup de travaux se sont intéressés à l'extraction sémantique par des caractéristiques signaux uni-modales dans le flux visuel, [Hamadi, 14] et audio, [Vinet, 13] ou texte, [Mallek, 11] et [Poignant, 11]. Ces travaux ont essayé de réduire l'écart relatif au passage du signal numérique à des descripteurs sémantiques, en raccordant le signal numérique par un ensemble de descripteurs de bas niveau liés aux modalités image, audio et texte.

Afin d'exploiter plusieurs sources d'informations hétérogènes de documents multimédia, de nombreux travaux se sont intéressés à la fusion des modalités (modalités visuelle composée d'images, textuelle composée d'un ensemble de termes et sonore), [Ercolessi *et al.*, 12]. Cette fusion peut avoir lieu à plusieurs niveaux du module d'indexation. Le contexte induit par l'usage de concepts intermédiaires et/ou de plusieurs sources d'informations fait aussi l'objet de travaux visant à combler le fossé sémantique pour une indexation sémantique de qualité, [Ercolessi *et al.*, 12] et [Bhatnagar *et al.*, 13].

Dans la suite de ce chapitre, nous commencerons par présenter les différents types d'indexation. Ensuite, nous citerons les différents travaux d'indexation des documents tout en décrivant les différentes structures des média, les différents niveaux sémantiques des descripteurs des média. Enfin nous mettrons en évidence le module de recherche d'information.

I. Catégories d'indexation

Trois catégories d'indexation existent dans la littérature : l'indexation manuelle, semi-automatique et automatique. Cette dernière est évidemment venue pour réduire la tâche de l'être humain, [Thiault, 11].

I.1. L'indexation manuelle

G. VAN SLYPE décompose l'indexation humaine en quatre étapes. La première est la prise de connaissance du contenu du document, puis le choix des concepts, ensuite la traduction des concepts en descripteurs et enfin l'établissement de liaisons entre les descripteurs, [Mustafa, 05].

L'indexation manuelle d'information est considérée coûteuse car pour indexer correctement un document, il faut faire intervenir des personnes qui soient elles-mêmes capables d'en comprendre, ce qui impose un coût non négligeable.

▪ Exemple

Le module d'indexation humaine d'un texte se déroule de la façon suivante : l'indexeur prend connaissance du document. Il lit rapidement le titre, la table des matières, le résumé, l'introduction générale, les introductions et les conclusions des principaux chapitres, les figures et la conclusion générale.

Les différents modèles d'indexation manuelle sont :

- l'indexation dite « à plat » où tous les descripteurs sont placés au même niveau d'importance ;
- l'indexation pondérée où nous distinguons les descripteurs principaux et secondaires, c'est la manière d'indiquer qu'un document traite prioritairement un sujet par rapport aux autres ;
- l'indexation à rôles où l'on souligne les relations qu'entretiennent entre eux les descripteurs retenus pour mettre en évidence de manière plus fine le sujet du document.

Le module d'indexation souffre de la subjectivité des descriptions vu que les indexeurs ne sont pas ceux qui envoient les requêtes. Ils demandent beaucoup de qualité de la part de l'indexeur.

I.2. L'indexation semi-automatique

Le choix final revient au spécialiste, qui intervient souvent pour choisir d'autres termes significatifs, l'indexation semi-automatique se divise en deux parties, une partie automatique permettant d'extraire une liste de descripteurs, et une deuxième partie qui est

manuelle réalisée par un spécialiste du domaine dont la tâche est de sélectionner des termes significatifs parmi les descripteurs retournés auparavant. Généralement le spécialiste utilise un vocabulaire contrôlé sous forme de thésaurus ou de base terminologique.

I.3. L'indexation automatique

L'indexation automatique est un domaine de l'informatique et des sciences de l'information premièrement apparu dans les bibliothèques. Elle utilise des méthodes logicielles pour établir un index pour un ensemble de documents et faciliter l'accès ultérieur aux documents et à leur contenu pour établir une liste de réponses rapprochées à la demande de l'utilisateur.

I.4. L'opposition entre l'indexation manuelle, semi-automatique et automatique

L'indexation automatique pourrait être le complément de l'indexation manuelle ou inversement, car l'indexation automatique nécessite un dialogue entre l'indexeur et l'ordinateur.

En conclusion, l'indexation automatique n'existe pas véritablement, c'est pour cela que l'on parle d'indexation assistée par ordinateur. En effet, on peut dire que les systèmes actuels s'ils remplacent l'homme pour une importante part de son expertise, ils ne le remplaceront pas complètement, car l'expression « indexation automatique » suppose une intervention totale du système, ce qui est loin d'être le cas puisque l'intervention humaine est toujours nécessaire.

En effet, avec ces progrès technologiques, l'enjeu est de diffuser aux utilisateurs des média de qualité dans une présentation cohérente, attractive et interactive, qu'elle soit destinée au monde de la télévision numérique, de l'internet ou aux personnes cherchant une formation particulière. Le but consiste à développer des technologies permettant la diffusion et la présentation de contenu multimédia dans des environnements aux contraintes multiples.

I.5. Objectifs

De nouveaux outils puissants et automatiques d'indexation et de recherche sont de ce fait vivement attendus dans de nombreux domaines allant de la vidéo sur demande à la télé-éducation en passant par les bases de données multimédia, les jeux, la surveillance, le commerce électronique, ainsi que les visioconférences dans les différents domaines. Tenons l'exemple de l'enseignement utilisant pour ses didacticiels le multimédia interactif : l'élève peut manipuler à l'écran les représentations des objets d'une expérience en science expérimentale, voir le résultat et éventuellement l'entendre. Dans cette thèse, nous étudions en particulier les visioconférences médicales. Ces dernières ont un caractère multimédia qui fait que la recherche par le contenu dans celles-ci présente un certain nombre de spécificités. Elles doivent leur richesse à la superposition sur un même sujet de l'image, du texte et du son. Par exemple, un concept donné (personne, objet) peut être présent avec différentes manières : il peut être visualisé, entendu (parole, musique et sons environnementaux), écrit textuellement (ressources textuelles, transcription des paroles et labels textuels qui décrivent le contenu). La combinaison de ces cas peut également se produire. Naturellement, ces distinctions sont

importantes pour l'utilisateur. Des requêtes impliquant un concept comme par exemple : « chercher les segments vidéos montrant une image de X » ou « chercher les segments vidéos dans lesquels on parle de X » sont susceptibles de produire des réponses différentes.

Dans le premier cas, on cherche X dans un flux visuel tandis que dans le second, on cherche dans le flux audio un segment dans la transcription duquel X est mentionné. L'automatisation de la phase d'indexation d'une visioconférence pose des problèmes très divers liés à la possibilité de confier une tâche souvent effectuée par l'être humain à une machine. En fait, l'homme est capable de distinguer et d'interpréter les différents aspects visuels, les relations entre les actions et la signification des aspects visuels et leur relation avec le contexte des actions et la classification des documents vidéos selon différents critères. Ceci se fait grâce à la possibilité de combiner les différentes informations acquises par le système visuel et le système auditif et la grande puissance d'analyse et d'interprétation chez l'homme.

L'objectif de cette thèse consiste donc à proposer un modèle de représentation du contenu sémantique des documents visioconférences permettant la prise en compte synthétique et intégrée d'éléments d'information issus de chacune des modalités (images, texte, son...) et de mettre en œuvre ce modèle dans le cadre d'un système de recherche par le contenu de documents visioconférences.

II. Types d'indexation

II.1. Indexation par concept

L'indexation des documents peut être effectuée par un groupe de concepts, où un document traite des concepts différents par exemple concept A et concept B sans savoir les relations entre ces concepts.

L'indexation automatique pose des problèmes notamment celui de l'ambiguïté des termes (homonymie et polysémie) fait recours à des outils de **T**raitement **A**utomatique des **L**angues (TAL). Mais ces techniques ne résolvent pas totalement le problème et il faut toujours faire le compromis entre la finesse des traitements et la complexité des systèmes. **Krovetz** a montré la nécessité d'indexer par les concepts (sens des mots) ainsi que les mots. L'indexation des documents par les concepts uniquement peut induire en erreur car les techniques de désambiguïsation ne sont pas absolument fiables, [**Krovetz, 97**]. Par conclusion, si nous nous basons uniquement sur ce type d'indexation nous risquons une perte d'information. La détermination du degré de similarité entre deux concepts reliés à des termes d'un document est un problème qui se pose dans beaucoup d'applications : désambiguïsation, résumé automatique, extraction d'information et indexation automatique, etc. [**Zargayouna et al., 04**].

- **Exemple d'indexation du document en concept**

Pour indexer un document médical, ce dernier va être découpé en phrases, chaque phrase est ensuite analysée pour en extraire les mots de référence. Nous effectuons une recherche des termes du thésaurus (défini dans la section II.1.3 du deuxième chapitre de l'état de l'art) qui

sont inclus dans la phrase c'est-à-dire que nous extrayons les termes dont tous les mots sont présents dans la phrase. Nous retiendrons ensuite les concepts représentés par ces termes. Et cela constituera l'indexation de la phrase.

Citons l'exemple situé par Sowa, dans la phrase "Néphrite glomérulaire lupique", nous reconnaitrons les concepts "Glomérulonéphrite lupique" et " Glomérulonéphrite" dans le **M**edical **S**ubject **H**eadings (MeSH).

Dans l'**A**ide au **D**iagnostic **M**édical (ADM), nous reconnaitrons les concepts "Néphrites glomérulaires" et "Lupus". Une option du module d'indexation permet de générer automatiquement les concepts pères des concepts extraits. Il est donc possible de générer les concepts "Glomérulonéphrite", "Néphrite", "Rein, maladies", (selon la relation "est_un" - "is_a" - de l'**U**nified **M**edical **L**anguage **S**ystem (UMLS)) à partir du concept "Glomérulonéphrite lupique".

Chaque concept étant décrit par un ou plusieurs termes. Ces concepts sont généralement organisés de manière hiérarchique (taxonomie, méronymie ...) et parfois de manière sémantique (réseau sémantique comme dans un graphe conceptuel, [Sowa, 84]).

Les théories d'indexation par concept existantes tentent d'apporter des éléments de compréhension pour définir les concepts, mais aussi expliquer leur structure et leur mécanisme d'acquisition. Nous distinguons trois théories.

- Selon la **théorie classique**, un concept est défini selon les caractères qu'il possède. Ces caractères sont nécessaires et suffisantes pour juger l'appartenance ou non d'un élément à une catégorie conceptuelle. Les caractères sont ceux qui font que l'élément est ce qu'il est, par exemple, dans le domaine de médecine, chaque maladie a ces caractéristiques qui se présentent sous forme de symptômes. Ces derniers, représentent une des manifestations d'une maladie, tel qu'exprimé par le patient. Les symptômes typiques de la grippe sont une fièvre élevée (au-dessus de 39°C), la fièvre survient rapidement et brusquement, associée à des frissons dus à l'augmentation rapide de la fièvre, des maux de tête importants, une fatigue extrême, des douleurs dans les articulations, des courbatures dans les membres (jambes par exemple), une toux sèche, des maux de gorge, un rhume et des nausées.

La classification des concepts a une structure arborescente, ce qui rend alors compte de la hiérarchie des concepts. Par exemple, les **M**aladies **H**éréditaires du **M**étabolisme (MHM), où un défaut dans le métabolisme intermédiaire affecte le foie, les muscles ou le cœur, ou encore le cerveau, et qui forment un groupe de maladies individuellement rares, sont inclus dans le concept maladies héréditaires donc toutes ces maladies sont des maladies héréditaires. Toutefois, les maladies héréditaires ne sont pas toutes des maladies du métabolisme, c'est à dire la réciproque n'est pas nécessairement vraie.

L'avantage est que cette théorie explique très bien ce qu'est un concept du fait des propriétés qui le caractérisent. Ainsi, le concept, selon la théorie classique, est une

définition de la forme suivante : x relève du concept si et seulement s'il possède les caractères c_1, \dots, c_n .

Pour conclure, la théorie classique est suffisamment vague pour n'utiliser qu'elle, étant donnée qu'elle ne définit pas explicitement la nature des caractéristiques qui définissent un concept.

- La **théorie empiriste** expose qu'à travers l'expérience, les sens et la perception sont à l'origine de l'acquisition des concepts, ce qui sous-entend que les caractéristiques seraient soit issues directement des sens, soit d'origine perceptive (les percepts).

Ainsi, il existe des concepts empiriques c'est-à-dire de concepts primitifs. L'apprentissage de nouveaux concepts relève soit de la composition de concepts, soit d'un processus plus complexe, appelé "inférence" pour dériver de nouveaux concepts à partir des éléments connus, [Ayache, 07].

En conséquence, la composition d'un concept n'est pas suffisante d'où l'apprentissage d'un concept concerne aussi l'apprentissage de son usage. La prise en compte de la sémantique du rôle conceptuel semble alors fondamentale. L'apprentissage du concept n'est donc qu'une relation de composition. Il concerne aussi les relations sémantiques entre des concepts (et pas seulement ceux qui le composent).

- La **théorie des rôles** conceptuels avance que les concepts ne sont pas seulement structurés hiérarchiquement comme la majorité des théories le suppose, mais aussi selon des groupes sémantiques. Les relations ne sont pas seulement des relations de composition mais elles sont aussi de nature sémantique. Ceci justifie la notion d'inférence ou de dérivation de concepts à partir d'autres concepts. Toutefois, d'après la théorie empiriste, cette théorie n'explique pas comment les premiers concepts se forment. Il convient alors de parler de concepts primitifs, [Ayache, 07].

Nous retenons plusieurs points issus de chacune des trois théories présentées :

- un concept est associé à un ensemble de caractéristiques (théorie classique) ;
- la perception joue un rôle fondamental dans la formation des concepts (théorie empiriste) ;
- les concepts sont reliés entre eux par des relations sémantiques (rôles conceptuels).

II.2. Indexation par contenu

Il existe plusieurs prototypes implémentant les systèmes d'indexation et de recherche du document multimédia par le contenu. Mais, jusqu'à présent, les systèmes de recherche par le contenu s'efforçaient de retrouver des documents ayant les mêmes caractéristiques que la requête pour ne pas retrouver uniquement des documents visuellement similaires et arriver à effectuer une réelle recherche sémantique.

Actuellement, plusieurs SRI s'intensifient. Et les nouveaux systèmes automatiques de recherche d'information par le contenu devront séparer le contenu visuel du contenu sémantique.

Dans ce qui suit, nous décrivons le principe d'indexation par contenu de chaque modalité ainsi que quelques travaux de recherche présentés dans la littérature, [Kharbouche, 06].

II.2.1. Indexation de texte

L'indexation d'un texte consiste à repérer dans son contenu certains mots ou expressions particulièrement significatifs (appelés termes d'indexation) dans un contexte donné, et à créer un lien entre ces termes et le texte d'origine. Il existe trois types d'indexation, [Poignant, 11] :

1. *manuelle* : lorsque le document est analysé par un spécialiste du domaine ou un documentaliste ;
2. *automatique* : lorsque le module d'indexation est complètement informatisé ;
3. *semi-automatique* : lorsqu'une première sélection de termes est réalisée automatiquement mais le choix final reste au spécialiste. Les systèmes les plus simples et les plus répandus sont basés sur la sélection de mots-clés dans les textes.

L'indexation manuelle assure une bonne correspondance entre les documents et les termes. Cependant, cette méthode demande un travail manuel qui est non seulement très difficile mais très long à réaliser par les indexeurs.

Salton, [Salton, 86] a démontré les inconvénients de ce type d'indexation, par exemple : deux indexeurs peuvent indexer deux documents identiques avec des termes différents et des différences d'indexation peuvent également exister chez la même personne qui indexe à des moments différents.

L'indexation automatique est complètement informatisée et est réalisée en plusieurs étapes : (i) l'extraction automatique des mots-clés qui correspondent au mieux au contenu informationnel du document, (ii) l'élimination des mots vides ou des mots fonctionnels (ex : conjonctions de coordination), (iii) la lemmatisation pour retrouver la racine des mots, (iv) la pondération des mots, pour affecter un poids élevé aux mots les plus importants.

Le résultat du module d'indexation par mots clés est un index plat, qui est une liste de mots alphabétiques et/ou thématiques sans aucune relation sémantique entre ses termes, les mots sont indépendants et les relations qui peuvent exister entre eux ne sont pas spécifiées.

L'indexation sémantique (cf. **Figure 1.1**) est une spécialisation de l'indexation classique, elle prend en compte la sémantique des mots au travers des relations entre les termes indexés.

Selon Baziz, [Baziz, 05], l'indexation sémantique consiste à utiliser les sens des mots afin d'indexer les documents en désambiguïsant les mots dans le document. Elle se distingue de l'indexation conceptuelle qui est, selon le même auteur, une généralisation de l'indexation sémantique et consiste à identifier des concepts dans les documents, ces concepts véhiculent du sens.

***Définition** : Nous définirons l'indexation sémantique comme l'utilisation d'ontologies de domaine ou de thésaurus afin d'effectuer le processus d'indexation, une ontologie étant un ensemble de concepts reliés par une*

relation de « spécialisation/généralisation » définissant un domaine donné. Chaque concept est dénoté par un ou plusieurs termes. [Baziz, 05].

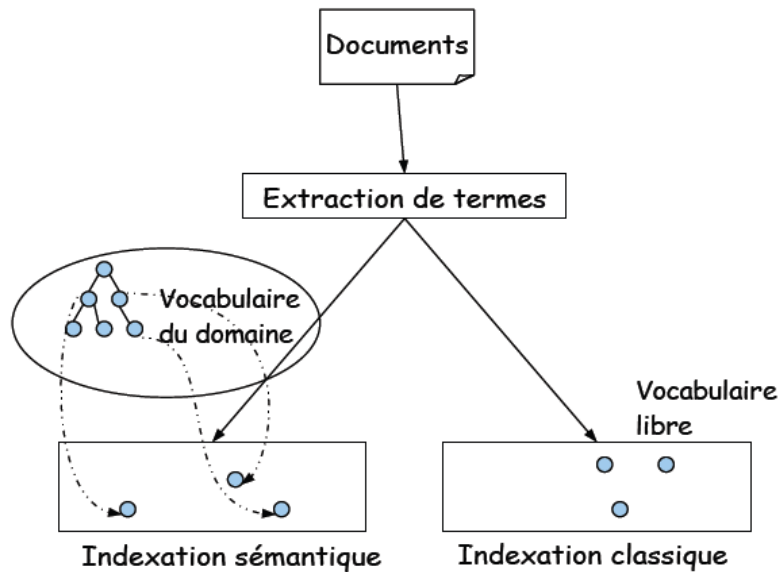


Figure 1. 1: Comparaison entre l'indexation classique et l'indexation sémantique

Il existe plusieurs travaux traitant l'utilisation de l'indexation sémantique. Nous en avons retenus trois que nous avons estimés pertinents afin de montrer l'utilité d'une telle approche :

- indexation avec une terminologie orientée ontologie : vu que l'information sur le Web n'est pas facile à retrouver, compte tenu de l'hétérogénéité et la mise à jour de cette information, une approche basée sur l'indexation de sites Web avec des mots-clés rattachés à des concepts les représentant a été proposée, [Abrouk, 06], [Gagnon, 13] et [Nedjar, 14] ;
- indexation dans le modèle *DocCore* : Ce modèle a pour but de représenter le contenu sémantique des documents en les projetant sur une ontologie linguistique générale. Contrairement aux réseaux sémantiques qui sont une structure de graphe avec des nœuds étiquetés par des constantes de relations, ce modèle construit un noyau sémantique pour chaque document avec les concepts et leur proximité. L'idée de cette approche est de représenter l'importance d'un terme dans un document avec : (i) sa fréquence d'apparition et (ii) la proximité avec les autres termes dans le document, [Baziz, 05] ;
- indexation avec une ontologie pour la désambiguïsation : Dinh, [Dinh et al., 10] utilise l'ontologie dans le module d'indexation afin de désambiguïser les termes extraits d'un document par la cooccurrence des mots-clés et la proximité sémantique des concepts dans l'ontologie. Dinh utilise un algorithme de désambiguïsation des concepts représentant un document et ceci par une liste de mots-clés.

Afin de réduire la taille d'un index et d'augmenter sa pertinence, nous excluons de l'index, toute une série de termes que nous appelons mots vides : nous retrouvons dans cette liste les articles, les pronoms, les conjonctions, les prépositions, certains adverbes, les déclinaisons afin de retrouver les racines des mots, etc.

Le choix des termes composant la liste des mots vides est très important et il est souvent préférable de laisser la possibilité à l'utilisateur de la base de données de modifier lui-même cette liste. En effet, par exemple, dans le cas, d'une base de données médicales, les termes « docteur », « hôpital », « traitement » seront tellement utilisés que nous retrouverons leur occurrence dans chaque enregistrement de la base de données : il conviendra donc d'incorporer ces termes à la liste des mots vides, [Ben Aouicha, 09].

Chaque unité documentaire peut alors faire l'objet d'une représentation vectorielle : les coordonnées représentent les fréquences des mots non vides. Lorsque nous effectuons cette opération pour un corpus de documents ou de pages web, nous obtenons une matrice dont les colonnes représentent un document et les coordonnées représentent la fréquence des termes permettant de repérer les mots les plus significatifs d'un texte (les descripteurs) afin que ceux-ci servent de clés d'accès.

Le flux textuel peut avoir occurrence non pas seulement dans un document textuel mais aussi dans un document vidéo. Dans ce cas, il contient une information riche en éléments sémantiques et relativement facile à exploiter de façon similaire aux documents textuels, [Poignant, 11].

II.2.2. Indexation des images fixes

Les systèmes d'indexation et de recherche des images décrits dans, [Attia, 13], [Cohen *et al.*, 10] et [Van De Sande *et al.*, 10] sont basés uniquement sur le contenu visuel extrait des images. Pour chaque image à indexer, le système calcule un index à l'aide des descripteurs extraits sur les trois attributs suivants : la couleur, la texture et la forme. Ces descripteurs caractérisent l'information en valorisant leurs proportions et leurs distributions relatives et absolues dans l'image. Ils font l'objet de plusieurs travaux de recherche dans le domaine de traitement des images et de la reconnaissance de formes (le choix de l'espace colorimétrique, les ondelettes, la quantification, le suivi de contours, ...). Les algorithmes utilisés dans ce cadre permettent de réduire le bruit et la redondance des informations avant de déterminer les points et les zones d'intérêts dans l'image, [LeCu *et al.*, 10] et [Krizhevsky *et al.*, 12]. À partir de ces informations, ils peuvent calculer le vecteur discriminant qui résume au mieux l'image dans la base. En phase de recherche, ces systèmes utilisent une image présentée par l'utilisateur comme une requête qui permet de rechercher et de présenter les images similaires de la base indexée. Une série d'échange entre l'utilisateur et le système (le bouclage de pertinence) permet d'affiner la requête. Malheureusement, ce type de système se limite aux concepts du contenu bas niveau car il est jugé incapable de saisir tous les concepts de haut niveau sémantique visés par l'utilisateur.

L'image requête est riche en informations, ce qui rend difficile la détermination de l'objectif et l'intention de l'utilisateur. Comme exemple, une image requête contenant un couche de

soleil sur la mer peut engendrer une ambiguïté sur le but de l'utilisateur : soleil, mer ou les deux. Pour faciliter le traitement de concepts de plus haut niveau, certains systèmes exploitent de manière conjointe les textes qui peuvent être codés avec l'image ou autour de l'image, [Ferecatu *et al.*, 05], [Florea *et al.*, 05] et [Philipp-Foliguet *et al.*, 06].

D'autres systèmes exploitent les textes figurant dans l'image par l'utilisation du processus de reconnaissance des caractères dans les images (par exemple www.riya.com). Ces derniers systèmes exploitent à la fois les concepts bas et haut niveaux de l'utilisateur et établissent les liens entre ces deux niveaux améliorant ainsi les performances. Cependant, ces techniques sont limitées par la difficulté d'indexer manuellement ou de façon semi-automatique de grandes quantités de documents, surtout pour les documents qui ne contiennent pas de données textuelles.

Dans [Ferecatu *et al.*, 05], les auteurs proposent une méthode qui combine des descripteurs visuels et des mots-clés. L'index visuel se détermine à partir de l'histogramme des couleurs, de la forme et de la texture en utilisant les transformées de « Hough » et de « Fourier ». En ce qui concerne les mots-clés, la méthode proposée décompose les mots-clés dans une base de données lexicale externe. La combinaison entre ces deux informations complémentaires se fait à l'aide de Machine à Vecteurs de Support (SVM¹) dont les données d'apprentissage sont extraites via le retour de pertinence. Les expériences menées avec cette approche sur une base de documents représentant des concepts liés aux activités humaines (art, archéologie, architecture, ...) montrent que la combinaison entre l'image et les mots-clés améliore significativement la précision du système.

Le système *Retimm* proposé dans [Philipp-Foliguet *et al.*, 06] a été conçu pour indexer et rechercher les images par leurs contenus et les mots-clés extraits des commentaires et des titres. Ce système accepte les requêtes multiples (images et/ou mots-clés) pour lesquelles le vecteur descriptif de l'image est calculé à partir :

- des distributions des couleurs (quantification sur 256 niveaux) et des textures (utilisation de filtre de Gabor) dans l'image ;
- d'un ensemble de mots-clés prédéfinis.

Le calcul de similarité ne s'effectue pas par une fonction de similarité ordinaire mais par un système de vote basé sur l'algorithme des *kplus proches voisins* dans la base d'apprentissage. Le calcul d'un score à valeur binaire 0 ou 1, comptabilisé pour chacune des caractéristiques du vecteur descriptif, permet d'éviter le calcul de la distance dans l'espace multidimensionnel. Le score final associé à chaque image de la base correspond à la somme des votes pour toutes les dimensions. La rapidité du système en phase de recherche est un avantage mis en évidence par les auteurs.

Nous citons aussi le modèle An Extended Model for Image Representation and Retrieval (EMIR²), [Mechkour, 95] et [Charhad, 05]. Ce dernier permet une représentation

¹Les Machines à Vecteurs Supports (SVM) constituent un algorithme de classification discriminatif qui a été introduit en 1995 par Vapnik pour la classification de texte et il est maintenant largement utilisé pour les applications vidéo.

symbolique du contenu sémantique des images fixes. Il décrit une image par les vues suivantes :

- **La facette perceptive**

La facette perceptive regroupe l'ensemble des attributs visuels associés à l'image entière mais aussi aux différents objets image tels que la couleur, la texture, la forme etc. Un des principaux avantages dans l'utilisation de ces attributs réside dans leur caractère objectif et dans la possibilité de les calculer de façon automatique.

- **La facette structurelle**

Une image est décomposée en objets images. Chaque objet image peut également être décomposé en sous-objets images. La vue structurelle de l'image a la forme d'un graphe orienté ayant pour nœud les objets de l'image et pour arcs des relations de composition structurelle. Elle permet de décomposer l'image en conservant la structure des éléments présents dans l'image.

- **La facette spatiale**

Cette facette permet de décrire les relations spatiales 2D dans le plan de l'image ou 3D dans la scène qui lie les images ensemble. Les objets sont positionnés les uns par rapport aux autres. La vue spatiale d'une image est un graphe dans lequel les nœuds sont des objets spéciaux et les arcs sont des relations spatiales.

- **La facette symbolique**

Cette facette associe une description sémantique à l'image et aux objets de l'image. La description regroupe un ensemble de termes pouvant être reliés ensemble par des relations sémantiques.

Généralement, un grand nombre de descripteurs numériques ont été proposés afin de décrire le contenu d'une image. Le plus souvent, les techniques d'extraction d'information liées à l'analyse d'une seule image, visent à segmenter cette image en régions afin d'en extraire des informations intermédiaires concernant les couleurs, les textures et les formes qu'elles contiennent.

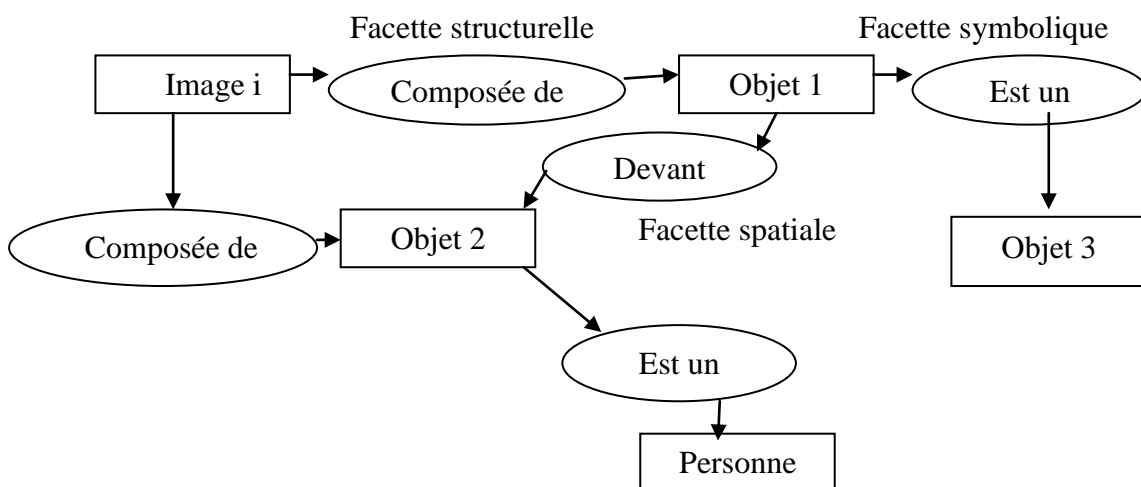


Figure 1. 2 : Facettes d'une image selon le modèle EMIR

II.2.3. Indexation de l'audio

Traditionnellement, l'indexation passe par une description textuelle de documents, cependant, appliquée aux contenus audio musicaux, l'association de mots-clés n'assure qu'une description limitée et généralement éloignée du contenu sonore (titre, compositeur, année, ...). Une description textuelle du contenu réel est néanmoins envisageable, par exemple en recensant les instruments utilisés dans un morceau de musique, ou bien en qualifiant un son particulier ...

L'indexation audio musicale a donc besoin d'outils semi-automatiques ou automatiques permettant une description objective du contenu sonore, afin d'assurer une gestion efficace des données.

La question de la description se pose en termes d'extraction automatique de l'information représentative du contenu sonore. Les descripteurs obtenus permettent de manipuler les documents selon leur contenu, en fonction de l'application visée.

Or, le document sonore est particulièrement difficile à indexer, car l'extraction de l'information élémentaire se heurte à l'extrême diversité des sources acoustiques. Il peut être intéressant de rechercher des « bruits » ou de sons sémantiquement significatifs tels que les applaudissements ou les effets spéciaux.

D'autres informations pertinentes peuvent être la détection de locuteurs signifiant toutes paroles dans un dialogue, ou encore leur identification s'ils sont connus à priori. La recherche de mots-clés (mots isolés, groupes de mots, ...) est une information importante sur le contenu du message verbal, [Bost *et al.*, 15].

Plusieurs travaux d'indexation sont présentés dans la littérature. En étudiant ces travaux, nous avons trouvé que l'indexation peut être faite en thèmes, en qualité (*bande large/bande étroite, bruitée/non bruitée*), en parole/non parole, en locuteurs, [Geoffroy, 13] et [Bost *et al.*, 15], etc... Dans le cas de la segmentation en locuteurs, le but est de découper le flux audio ou segments précisant pour chacun le début, la fin ainsi que l'étiquette du locuteur auquel il correspond.

La plupart des systèmes de segmentation, [Dessein *et al.*, 11] est composé par les étapes suivantes :

- para-métrisation du signal de parole (calcul de paramètres acoustiques) ;
- pré-segmentation acoustique (l'enregistrement est découpé en segments de parole, silence, musique, etc.) ;
- détection des changements du locuteur ;
- regroupement des segments (les segments trouvés pendant la phase précédente sont regroupés et le nombre de locuteurs est estimé).

Le système automatique de segmentation en locuteurs repose sur la détection des changements des locuteurs au sein d'un fichier audio associé à un processus de regroupement de tous les segments audio appartenant à un même locuteur. La sortie de SLR répond à la question « qui parle et quand ? » [Béchet *et al.*, 15]. La tâche SLR peut être à différents types

de documents sonores : conversations téléphoniques, émissions radiophoniques, enregistrements de réunions.

II.2.4. Indexation de la vidéo

L'accès aux contenus des vidéos peut être scindé en deux grandes familles : les contenus sémantiques et les contenus automatiques des vidéos. Les contenus sémantiques sont les descriptions d'une vidéo faite par un utilisateur, par exemple des mots-clés, des résumés de séquence, etc. Les contenus automatiques se réfèrent à l'ensemble des descripteurs qui peuvent être extraits via la vidéo elle-même. Beaucoup de ces descripteurs sont communs avec les descripteurs automatiques du contenu d'une image. De nombreux travaux de recherche ont abordé le problème de recherche par le contenu d'une image, plusieurs états de l'art de ces méthodes existent dans [Ribeiro *et al.*, 13]. Les vidéos apportent une dimension temporelle par rapport aux images. De nombreux descripteurs automatiques sont alors ajoutés tels que la hiérarchisation en séquences, plans et frames ou encore la notion de mouvement des objets ou de la caméra, [Farabet *et al.*, 13], [Airouche, 12] et [Rombaut *et al.*, 10]. Les vidéos comme la musique ont une dimension sonore : par exemple, Zhang *et al.* [Zhang *et al.*, 06] se sont basés sur les descripteurs automatiques audio pour mettre en valeur des séquences vidéo sur le sport. La majorité de ces descripteurs automatiques du contenu des vidéos sont inclus dans la norme Mpeg-7, norme de description des contenus multimédia et audiovisuels utilisant le langage XML. L'ensemble de ces descriptions est stocké dans une base de fichiers annexes aux vidéos.

Dans [Zhou *et al.*, 07], les auteurs proposent une amélioration de leurs travaux antérieurs qui ne prenaient pas en compte la dimension temporelle apportée par la vidéo. L'idée principale est de décrire les vidéos de la base d'une manière plus compacte que par l'ensemble des caractéristiques de chaque image composant la vidéo. Pour cela, ils proposent de regrouper les images similaires d'une vidéo en plusieurs classes. Puis, la vidéo est reconstruite en remplaçant chaque image par identifiant de la classe d'image. Ainsi, une vidéo est décrite par une chaîne ordonnée d'identifiants de classe, ce qui permet de garder la dimension temporelle. Ben Abdelali *et al.* [Ben Abdelali *et al.*, 09] utilisent les descripteurs dynamiques locaux des vidéos pour indexer leur moteur de recherche interactif. Les descripteurs dynamiques locaux permettent d'estimer les mouvements des caméras et les mouvements d'objets dans la séquence vidéo à partir des points caractéristiques des séquences d'images formant la vidéo. Ayache *et al.* [Ayache *et al.*, 06] présentent leur système de détection des frontières des séquences vidéo basée sur les différences entre les images voisines et les descripteurs de mouvement. Ils proposent aussi une catégorisation des séquences vidéo par les SVM à partir de descripteurs hauts niveaux des séquences vidéo. Ces descripteurs hauts niveaux sont obtenus par la combinaison et la fusion des descripteurs bas niveaux (couleur, texture), [Attia, 13] et [Cohen *et al.*, 10], des données audio, des données textuelles et des contextes conceptuels. Ils exploitent leurs propositions dans un moteur de recherche vidéo par le contenu.

Nous avons proposé dans cette section un aperçu de quelques méthodes d'indexation et de recherche liées aux contenus vidéo. La plupart de ces méthodes sont basées sur une représentation des vidéos comme une suite d'images contenant de nombreux descripteurs extraits automatiquement.

II.3. Indexation par le contexte

Les systèmes d'indexation multimédia basés sur le contexte sont souvent des systèmes qui ont été mis en œuvre pour traiter les photos et les vidéos prises par des appareils cellulaires (téléphone mobile, PDA, ...). Contrairement à l'indexation basée sur le contenu, seules les informations issues du contexte dans lequel se trouve le document lors de sa création sont prises en considération. Dans de nombreux cas, les informations contextuelles semblent suffisantes pour satisfaire l'attente de l'utilisateur désireux de rechercher ou d'indexer des documents. Les systèmes peuvent être différenciés entre eux par le type d'information contextuelle qu'ils utilisent pour l'indexation, [Kharbouche, 06] et [Mothe, 11].

Ainsi, dans [Lajmi, 11], les auteurs utilisent l'instant et le lieu où l'image a été capturée. Ces informations sont utilisées pour ajouter un certain nombre de métadonnées associées à chacune des images. Plusieurs catégories de métadonnées peuvent être prises en compte : le nom du lieu (en décomposant le pays, la ville, la rue, le parc, ...), la saison, la période dans la journée (matin tôt, matin, après-midi, soir, nuit,...), la météo (pluie, ciel dégagé, beau ...), la température et le décalage horaire. Ainsi les photos sont organisées et annotées selon les métadonnées produites par le système.

Sur la même idée, nous trouvons une organisation des photos selon deux structures hiérarchiques : une hiérarchie des événements et une hiérarchie des lieux, [Farabet *et al.*, 13]. Ces deux structures hiérarchiques sont construites à l'aide de règles de discrimination exploitant les informations de date et de lieu. Les structures interfèrent entre elles dans la mesure où un lieu peut contenir plusieurs événements alors qu'un événement ne contient qu'un seul lieu.

À partir d'informations contextuelles (dates et lieux) et d'une série d'images complètement ou partiellement annotées, le système décrit dans [Lajmi, 11] permet de retrouver l'identité des personnes qui se trouvent dans une image. Ainsi, pour chaque image, il propose un ensemble de personnes candidates à partir d'un ensemble d'estimateurs : un estimateur de temps, un estimateur d'événement, un estimateur de lieu, un estimateur de lieux voisins, un estimateur de temps proche et un estimateur de coprésence. Pour une image à indexer, chaque estimateur attribue un score, qui prend la forme de probabilités estimées à partir de statistiques sur les images déjà annotées, à chaque personne d'un groupe prédéfini par l'utilisateur. Finalement, un estimateur global, présentant les candidats à l'utilisateur, réalise la fusion des autres estimateurs via deux méthodes distinctes. La première, appelée *Padding*, consiste à sélectionner dans chacune des listes ordonnées générées par les différents estimateurs les personnes candidates suivant leurs fiabilités respectives. La deuxième méthode fusionne les scores des différents estimateurs par une combinaison linéaire pondérée. Les expériences

réalisées sur ce système montrent que la combinaison linéaire donne les meilleures précisions, [Hamadi, 14].

L'indexation des photos se fait par l'analyse des dates de création qui sont regroupées par des algorithmes de classification tels que les *K-means* ou la classification hiérarchique descendante.

L'une des applications les plus abouties dans ce domaine est sans doute le système mis en œuvre pour faciliter le partage des photos prises par des téléphones cellulaires décrits dans [Davis *et al.*, 06]. Pour chaque image capturée, on s'intéresse à la création d'une liste, la plus pertinente possible, des éventuels destinataires de cette image. Cette liste est estimée à partir du traitement de trois types d'informations contextuelles : des données temporelles, des données spatiales et des informations de nature sociale, [Bermejo *et al.*, 11]. Les informations temporelles (date de capture) et spatiales (identification de la cellule radio) sont issues des équipements de réseaux mobiles alors que les informations sociales (identité du propriétaire, coprésence d'un autre utilisateur au même endroit, etc.) sont issues de serveur web MMM2 qui collectent et traitent toutes les données disponibles. Une adaptation pour passer du contexte au contenu a été étudiée en identifiant le nom du lieu où l'image a été prise. Ceci est rendu possible grâce aux images déjà annotées de la même cellule radio par d'autres utilisateurs et aux informations contextuelles extraites. Pour fusionner toutes ces informations, ce système utilise la combinaison linéaire avec des coefficients préfixés sans traiter les informations visuelles présentes dans l'image. Dans [Davis *et al.*, 06], on utilise des caméras équipées de GPS et d'une interface Bluetooth. L'équipement GPS permet d'identifier les coordonnées exactes de l'appareil photo contrairement aux coordonnées géographiques qui laissent une marge d'erreur de plusieurs mètres à plusieurs kilomètres selon la taille de la zone de couverture radio.

II.4. Indexation par le contenu et le contexte

Malgré la cohérence qui peut exister entre le contexte et le contenu d'un document multimédia, nous ne trouvons, dans la littérature, que peu de travaux qui manipulent simultanément ces deux types d'informations qui peuvent être complémentaires. Dans [Lajmi, 11], les auteurs n'exploitent que des informations extraites du contexte (les informations spatio-temporelles associées à l'image) pour identifier les identités des personnes figurant dans l'image à indexer. Autrement dit, nous remontons au contenu du document via son contexte. Les auteurs rejettent même parfois l'utilisation d'outils d'indexation basés sur le contenu (détecteurs de visages par exemple) pour identifier le nombre de personnes présentes dans l'image. Ils justifient ce choix par le manque de fiabilité et de sensibilité de ces dispositifs qui apporteront de nombreuses fausses détections ou des non-détections, dégradant ainsi la performance du système.

Dans [Davis *et al.*, 06], des descripteurs du contenu sont utilisés pour identifier les lieux et les visages dans les images prises par des téléphones cellulaires. Pour identifier les lieux, le système se base sur les informations contextuelles issues du système MMM2. L'analyse de l'image est effectuée par la méthode Cognitive Visual Attention (CVA) et la fusion se fait par

une combinaison linéaire pondérée dont l'ajustement est réalisé par une base d'apprentissage. Pour identifier les visages, plusieurs algorithmes de reconnaissance de visages sont mis en compétition.

III. Fossé sémantique et fusion multimodale

Si nous savons parfaitement interpréter et détecter le contenu d'une image ou d'une information, nous trouvons que le traitement d'une grande quantité d'informations est limité. Un comportement inverse de celui de la machine, qui ne trouve pas de problème à traiter une tâche répétitive donnée, mais limitée lorsqu'il s'agit d'interpréter automatiquement de données numériques.

Le fossé sémantique se réfère ainsi au manque de coïncidence entre l'information qui peut être extraite à partir d'une donnée visuelle et l'interprétation de cette même donnée pour un utilisateur dans une situation spécifique. Il sépare les représentations visuelles brutes (bas-niveau) et conceptuelles (haut-niveau).

Pour limiter les conséquences de cette problématique, Rachid Benmokhtar introduit, dans sa thèse Fusion multi-niveaux pour l'indexation et la recherche multimédia par le contenu sémantique 2009, le système **C**ontent **B**ased **I**nformation **R**etrieval (CBIR) basé sur le contenu et tenant compte de différents types de descripteurs et les avancées scientifiques dans le domaine de l'apprentissage automatique et de la fusion multi-niveaux. Ce système est utilisé dans le but de combiner des informations hétérogènes issues de plusieurs sources pour obtenir une information globale, plus complète et de meilleure qualité, [Benmokhtar, 09]. Une indexation efficace nécessite par conséquent une approche multimodale dans laquelle la collaboration de différentes modalités est utilisée. L'intégration des trois modalités est courante dans le cadre de l'interprétation sémantique des informations de la vidéo par des indexeurs humains, [Lu *et al.*, 10], [Feng *et al.*, 12] et [Ercolessi, 13].

Ainsi, l'indexation de la vidéo par le contenu est une récente application multimédia de grande importance actuelle. Dans ce cadre des efforts très significatifs ont été consacrés. Pour cette raison, des énormes budgets ont été consacrés ces dernières années aux projets de recherche dans l'indexation multimédia et la création de moteur de recherche (*cf.* **Tableau 1.1**).

Nous avons eu plusieurs modèles d'indexation des documents audiovisuels basés sur le contenu sémantique. À titre d'exemple Mbarek Charhad a présenté en 2005, [Charhad, 05] une extension du Modèle EMIR² pour l'indexation et la recherche dans une base de documents vidéo. Cette extension a nécessité la définition des nouveaux concepts et de nouvelles relations permettant de décrire le contenu du document vidéo.

L'approche de Mbarek Charhad consiste à considérer chaque scène vidéo comme une combinaison des diverses interprétations du contenu. Ces interprétations se traduisent par un ensemble d'objets vidéo et de relations qui les lient. Chaque interprétation est alors appelée « vue » ou bien « facette » qui peut être composée d'autres vues. Les vues sont classées selon l'interpréteur et selon les objets décrits.

Tableau 1. 1: exemples de projets de recherche dans l'indexation multimédia et la création de moteur de recherche, [Benmokhtar, 10]

Projet	Durée	Budget Euro	Financement
DELOS (Network of Excellence on Digital Libraries)	2004 à 2007	15.7	Union européenne
MUSCL (Multimedia Understanding through Semantics, Computation and Learning)	2004 à 2008	6.9	Union européenne
K-SPACE (Knowledge Space of Semantic Inference for Automatic Indexation and Retrieval of Multimedia Content)	2006 à 2008	8.9	Union européenne
THESEUS	2007 à 2011	120	Allemagne
Europeana	2008 à 2010	120	Union européenne
QUAERO	2008 à 2013	250	France

La description symbolique du contenu vidéo est une description multimodale. En effet, un concept peut être issu de la description du contenu visuel, du contenu audio ou bien du contenu texte.

Il a mis en œuvre des formes de représentation :

- Générique : cette forme de représentation regroupe l'ensemble des facettes décrivant les caractéristiques communes dans le document vidéo indépendamment du type de média, telles que par exemple, la nature temporelle de la vidéo. Nous distinguons deux types de facettes :
 - o Facette temporelle : c'est l'ensemble des relations temporelles qui relient les éléments d'informations dans le document vidéo.
 - o Facette événementielle : pour la description des différents événements contenus dans un document vidéo. Un événement est une ou plusieurs actions.
- Spécifique : permet de décrire le contenu vidéo par média tenant en compte les spécificités de la vidéo qui sont propres à un médium particulier (visuel, audio ou texte). Cette représentation contient :

- Facette sémantique : permet d'associer une description sémantique au contenu visuel, audio ou textuel. Cette facette est composée de trois autres facettes :
 - une sous-facette visuelle pour décrire le contenu visuel ;
 - une sous-facette audio pour la description du contenu audio ;
 - une sous-facette texte pour interpréter les informations textuelles.
- Facette signal : elle permet de décrire les caractéristiques de bas niveau pour des descriptions sémantiques. Elle regroupe plusieurs sous-facettes :
 - une sous-facette couleur pour représenter les caractéristiques couleurs du contenu visuel. Par exemple, les couleurs attribuées à une image x ;
 - une sous-facette texture décrivant les propriétés texture du contenu visuel ;
 - une sous-facette spatiale spécifiant les relations spatiales des objets images ;
 - une sous-facette mouvement définissant les mouvements de la caméra des objets images ainsi que leurs trajectoires.
- Facette événementielle : elle permet de décrire les événements sachant qu'un événement est un fait qui survient à un moment donné. Il signifie généralement tout ce qui arrive et possède un caractère spécifique. Dans le cas de la vidéo, un événement est quelque chose qui arrive dans une vidéo ou dans un segment vidéo (interaction entre les entités audiovisuelles (personne, objets, etc.)).
- Facette temporelle : c'est une caractéristique spécifique du document vidéo qui permet de synchroniser et d'ordonner les descriptions utilisées pour la modélisation du contenu vidéo et donne au document son caractère dynamique. Cette facette contient l'ensemble des relations qui relient les entités d'une vidéo dans un ordre temporel bien déterminé.

En conclusion, nous pouvons dire qu'il existe trois catégories d'indexation des documents :

- Une indexation à niveau haut, dite aussi sémantique, employant un ensemble de termes ou bien de concepts, fournit une description de haut niveau de ce que contient une vidéo. Il s'agit du temps, des lieux, des personnages et de leurs interactions... etc.
- Une indexation basse niveau qui permet d'accéder au contenu des vidéos par des descripteurs de bas niveau simple (la couleur, la texture...etc.). Les algorithmes d'analyse d'un signal extraient ces descripteurs, les organisent et emploient les techniques de recherche par similarité pour la recherche de séquences vidéos.

- Une indexation spécifique à un domaine dont ces techniques emploient la structure de haut niveau de la vidéo pour contraindre l'extraction des descripteurs visuels de bas niveau. Elles sont efficaces seulement dans leurs domaines prévus.

IV. Les descripteurs

IV.1. Les descripteurs de niveau bas

Les descripteurs de niveau bas sont des éléments d'indexation qui peuvent être extraits automatiquement par application d'algorithmes et sans connaissance particulière du contexte. Ils expriment des propriétés significatives directement observables ou calculables à partir du signal. Ils s'expriment sous forme numérique et sont stockés au sein de vecteur dont les attributs correspondent aux différentes caractéristiques du descripteur, [Bertrand, 06].

IV.1.1. Descripteurs image de bas niveau

Un grand nombre de descripteurs a été proposé afin de décrire le contenu d'une image. Les plus souvent utilisés, ceux qui sont relatifs aux techniques d'extraction d'information liées à l'analyse d'une seule image, visent à la segmenter en régions afin d'en extraire des informations intermédiaires concernant les couleurs, les textures et les formes qu'elles contiennent, [Van De Sande *et al.*, 10].

Le standard MPEG-7 (**M**oving **P**icture **E**xperts **G**roup) propose, dans sa partie dédiée à la vidéo, des schémas de description qui permettent la description des couleurs des régions d'une image (ou d'une image entière, ou d'un groupe d'images) par détermination de différents espaces de couleur, en utilisant les couleurs dominantes ou des histogrammes.

IV.1.2. Descripteurs audio de bas niveau

Au niveau local, les descripteurs numériques sont calculés sur un segment d'un échantillon à quelques dixièmes de secondes. Nous distinguons plusieurs familles de descripteurs liés au traitement appliqué pour les extraire :

- descripteurs temporels : ils sont extraits à partir de la forme d'onde ou de la courbe d'énergie (enveloppe). Nous trouvons le volume, caractéristique la plus largement utilisé. Simple à calculer, le volume est un très bon indicateur de silence, utile pour la segmentation audio. Le taux de passage par zéro est aussi très largement utilisé, il est notamment très efficace pour discriminer les dialogues de la musique. Typiquement, la voix est caractérisée par un faible volume et un taux important de passage par zéro. Enfin, le pitch, fréquence fondamentale du signal audio, est un paramètre important pour l'analyse et la synthèse de la voix et de la musique. Normalement, seules la voix et la musique harmonique ont des pitch caractéristiques ;
- descripteurs d'énergie : ils décrivent l'énergie de différents types de contenus : énergie globale, énergie harmonique, énergie du bruit, énergie de bandes, [Codreanu, 10] ;

- descripteurs spectraux : ils sont calculés à partir d'une représentation 2D du signal (fréquence, temps), à l'aide de la transformée de Fourier à court terme (STFT) ou de la transformée en ondelettes (DWT) ;
- descripteurs harmoniques : ils sont calculés à partir d'un modèle harmonique du signal ;
- descripteurs perceptifs : sont les plus utilisés pour la caractérisation et la reconnaissance de locuteurs. Ils sont déterminés à partir d'un modèle d'écoute humaine : Loudness, Sharpness, "**M**el-**F**requency **C**epstral **C**oefficients" ou MFCC, [Codreanu, 10]. Les MFCC se présentent sous la forme d'une représentation 2D (fréquence, temps). L'échelle des fréquences est fixée afin de correspondre à la capacité cognitive de l'oreille humaine. Les MFCC ont l'avantage de considérer la propriété non-linéaire de la perception humaine des sons selon plusieurs fréquences.

IV.1.3. Descripteurs vidéo de bas niveau

Les descripteurs numériques de la vidéo correspondent généralement à des interprétations en termes de couleur, de texture et de forme. Ces informations résultent de l'analyse de chaque image ou de segments d'images de la vidéo, [Attia, 13], [Cohen *et al.*, 10].

De nombreuses approches réduisent le problème de l'extraction du contenu d'un segment d'images au traitement des descripteurs d'une seule image du segment considéré. Plusieurs algorithmes ont été développés afin de déterminer l'image « moyenne » ou résumée d'une séquence d'images. C'est ce qui est fait dans [Tapu, 12], [Chesnais *et al.*, 12] et [Huart, 07] pour segmenter la vidéo en scènes.

D'autres approches proposent de définir et d'utiliser le mouvement d'objets visibles et/ou les mouvements de la caméra pour l'indexation de la vidéo, [Airouche, 12], [Rombaut *et al.*, 10] et [Delezoide, 06]. La méthode commence par une segmentation basée sur la couleur de l'image. Puis les régions sont regroupées selon leurs paramètres de mouvement.

IV.2. Les descripteurs de niveau moyen

Au niveau intermédiaire (appelé aussi niveau moyen) et global (appelé aussi niveau haut), les descripteurs sont issus de statistiques et d'opérations classiques sur les descripteurs (bas, intermédiaire et global) : moments statistiques, maximum, minimum, médiane, dérivées, corrélations, valeur de modulations, entropie, [Codreanu, 10].

Les descripteurs de niveau moyen expriment des propriétés ayant une signification plus évidente, i.e. plus compréhensible, par un utilisateur, comme le nombre d'instruments dans un morceau de musique. Les derniers efforts convergent vers l'extraction de concepts. Ces derniers sont dits qualitatifs lorsqu'ils s'expriment par une description naturelle du langage. Les concepts sont dits quantitatifs lorsqu'ils s'expriment par des nombres, comme le nombre de personnes dans une image, [Bertrand, 06].

À un concept est associée une taxonomie : l'ensemble des valeurs qu'il peut prendre. La valeur prise par un concept pour un document donné est déduite par classification automatique ou indexation manuelle du document selon la taxonomie choisie.

IV.2.1. Descripteurs image de niveau sémantique moyen

Le niveau sémantique moyen correspond aux concepts simples extraits du contenu de l'image. Les descripteurs moyens pour l'image sont en général de granularité spatiale intermédiaire. Ils sont déduits par catégorisation automatique de zones d'images dans des classes prédéfinies. Ces zones sont le plus souvent issues de la segmentation spatiale d'une image et sont homogènes vis à vis des caractéristiques liées à la classification choisie. La tâche de classification nécessite le calcul de modèles numériques complexes appris à partir de zones d'images annotées. Ces descripteurs concernent la détection de figures particulières :

- Objets : voiture, visage, ...
- Textures : herbe, arbre, eau, ciel, bâtiment, neige, sable, ...

IV.2.2. Descripteurs audio de niveau sémantique moyen

Le niveau sémantique moyen correspond aux concepts simples déduits du contenu sémantique bas audio. Ces concepts sont déduits par catégorisation automatique de segments de son dans des classes prédéfinies.

Plusieurs descripteurs moyens ont été proposés afin de décrire les caractéristiques d'un son. Ils concernent des informations variées : le type de la source d'émission (parole/musique/bruit), [Vincent, 12], [Parvaix, 10] et [Delzoide, 06], le nombre de locuteurs dans un dialogue, [Bost *et al.*, 15], la présence du silence, [Moussallam, 12].

IV.2.3. Descripteurs vidéo de niveau sémantique moyen

La description sémantique de niveau moyen d'une vidéo s'appuie, comme pour l'audio et l'image, sur la notion de concept. Un concept représente une description symbolique de la vidéo ou d'une partie de la vidéo.

Le système vidéo développé par IBM utilise un lexique varié pour représenter des concepts moyens audio (*musique/dialogue/monologue*) et image (Objets : *visage, voiture, bâtiment, pont, ...*, Textures : *ciel, eau, désert, neige, ...*). Aujourd'hui, l'amélioration des performances de l'indexation audio et image autorise la détermination automatique d'un grand nombre de concepts simples, [Ben Abdelali *et al.*, 09] et [Béchet *et al.*, 12].

Plusieurs descripteurs de niveau moyen peuvent être extraits de façon automatique à partir d'une **image** « moyenne » d'un segment de vidéo. Pour le **son**, les concepts moyens de granularité locale sont extraits à partir de segments d'audio de taille variable.

IV.3. Les descripteurs de niveau haut

Les descripteurs de niveau haut (appelé aussi global) expriment des propriétés qui ont une signification forte. Ce sont des concepts qui peuvent être décrits en termes d'autres concepts. Leur catégorisation automatique nécessite le calcul de modèles d'apprentissage textuels et numériques à partir de descripteurs de niveaux bas et moyen issus de documents

d'exemples. Dans le cadre d'analyse de films, ils expriment trois grands types de caractéristiques : le temps, le lieu et les relations entre objets (ou personnes) : actions, [Ben Abdelali *et al.*, 09] et [Béchet *et al.*, 12].

IV.3.1. Descripteurs image de niveau sémantique haut

Le niveau sémantique haut correspond à des concepts complexes que l'on extrait d'une image. Ces concepts sont déduits par classification d'une zone (ou de l'intégrité) d'une image dans des classes prédéfinies. Plusieurs descripteurs de ce type ont été développés dans la littérature.

Il s'agit principalement de la reconnaissance du lieu où une photo a été prise et de l'identification d'une personne à partir de son visage. Les concepts issus de la classification d'ambiance sont de granularité globale. Les concepts issus de la reconnaissance de personnages sont de granularité intermédiaire. Ces tâches de classification nécessitent le calcul de modèles numériques complexes appris à partir d'images de référence, [LeCu *et al.*, 10] et [Krizhevsky *et al.*, 12].

IV.3.2. Descripteurs audio de niveau sémantique haut

Le niveau sémantique haut correspond aux concepts complexes déduits par classification des descripteurs bas et moyen de segments de sons. L'automatisation de l'indexation pour ce type de descripteurs est toujours un sujet de recherche. Nous citons notamment la reconnaissance du locuteur, [Roy, 11], l'identification d'un style de musique, [Camus, 11] ou du nom d'un morceau, [Sargent *et al.*, 10].

IV.3.3. Descripteurs vidéo de niveau sémantique haut

Le niveau sémantique haut correspond à des concepts complexes que l'on extrait de la vidéo. La détermination de concepts est améliorée par la fusion de plusieurs média.

La reconnaissance de personnes est un excellent exemple de la nécessité de l'intégration multimodale. Identifier des personnages uniquement par le signal visuel n'est pas efficace à cause de la variance forte de l'orientation, l'illumination et l'occultation des visages au sein d'une scène. De plus, les détections de parole et l'identification du locuteur sont très sensibles aux bruits environnementaux. La combinaison de ces deux types d'informations peut être utilisée pour améliorer les performances générales de la détection des personnes.

De même, plusieurs concepts liés à la reconnaissance de lieu ou de localisation (*intérieur/extérieur, parking/magasin/maison, ville/paysage*) sont extraits pour les modalités visuelles et auditives. Le challenge, pour améliorer la localisation pour la vidéo, est donc de réaliser la fusion de l'information apportée par les deux média, visuel et auditif, afin que les deux collaborent à l'extraction des concepts considérés.

V. Module de recherche d'information

Pour faciliter l'accès à l'information ainsi que son exploitation au mieux, nous avons besoin de décrire cette information. Dans notre contexte, ceci se traduit par des documents annotés par un ensemble de termes représentatifs suivants ou non un vocabulaire contrôlé, et ceci de manière automatique ou bien guidée par des experts à différents degrés. Nous

illustrons dans la **figure 1.3** le module de RI et par cela l'étape d'indexation. Le problème central en RI est de localiser les termes que d'autres ont utilisés dans le document recherché, [Blair, 90].

Définition : Nous définissons un document comme le support d'une information.

Le module de recherche d'information se compose de deux phases principales :

- la phase d'indexation : qui a pour but de représenter au mieux le contenu du document ;
- la phase de recherche : qui consiste à restituer à un utilisateur les réponses les plus pertinentes par rapport à sa requête, en utilisant l'indexation des documents.

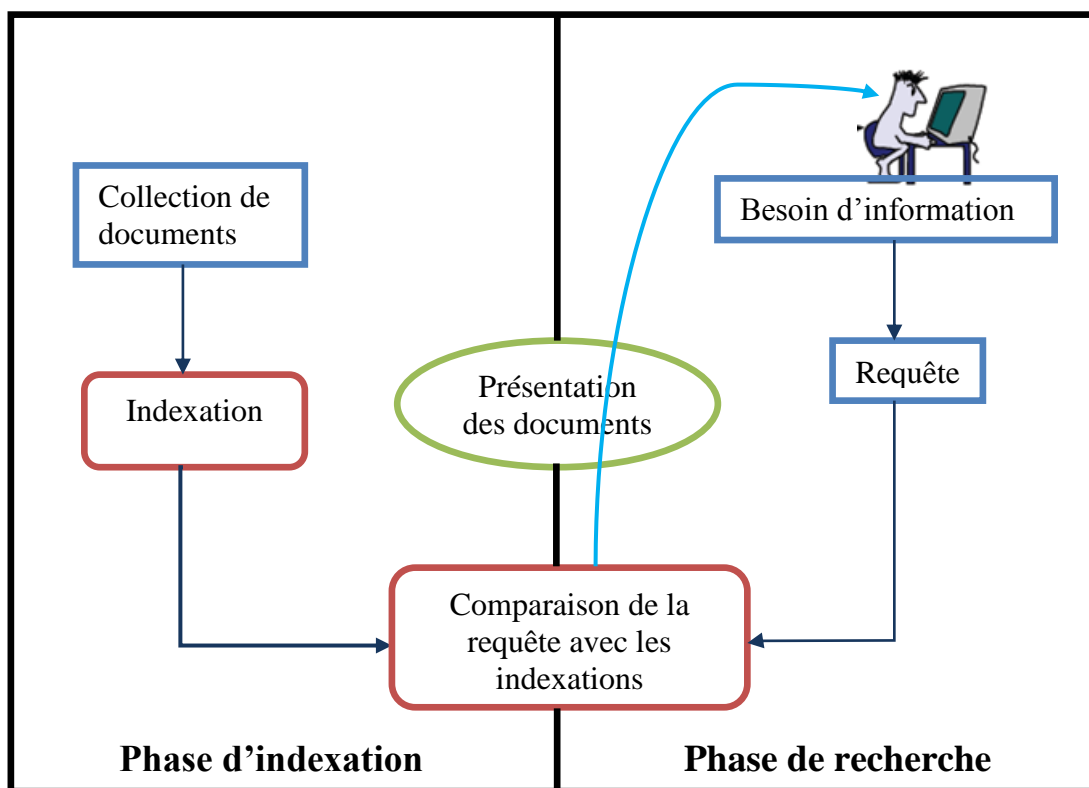


Figure 1.3 : Les étapes dans le module de recherche d'information

Conclusion

Dans ce chapitre, nous avons présenté un panorama des travaux liés aux problèmes de l'accès aux documents audiovisuels (DAV) et de leur manipulation. Plus précisément, nous nous sommes intéressés aux travaux sur l'indexation et la recherche par le contenu. Nous avons aussi présenté les aspects liés à la structure du contenu audiovisuel et les outils et schémas de description.

Le problème majeur lié au développement d'un système d'indexation et de recherche des DAV dans un contexte générique consiste à la mise en place des techniques capables de

combiner les sous-média pour traiter les documents. En effet, l'application des études spécifiques à un type de média (image, texte ou audio) telles que proposées dans la littérature est loin d'être efficace pour une application sur un DAV. Concernant l'utilisation de l'ensemble des outils et des schémas de description vidéo, ces outils ne permettent pas de décrire le contenu et les besoins du document vidéo d'une manière à tenir en compte à la fois de la richesse du contenu (contenu audio, visuel, etc.) et de la diversité des utilisateurs avec leurs besoins et leurs différents points de vue. Pour indexer le contenu de haut niveau dans un contexte général, l'intervention de l'opérateur humain reste à l'heure actuelle indispensable. Le résultat d'une indexation manuelle ou semi-automatique du contenu dépend directement du niveau d'expertise de l'opérateur et du temps qui lui est alloué pour associer une description à chaque segment vidéo. Généralement, les index se présentent sous une forme textuelle qui peut contenir soit des informations génériques (nom, type de fichier, titre, taille, ...) introduites lors de l'acquisition d'un DAV, soit des informations sur la structure du document comme par exemple la structure hiérarchique (séquence, scène, plan).

La notion de description sémantique ne coïncide généralement pas avec les méthodes d'analyse des images ou de l'audio. Ceci n'exclut pas l'existence de plusieurs approches « mixtes » qui intègrent les descriptions niveau signal (couleur, texture, mouvement de caméra, etc....) générée automatiquement avec les descriptions haut niveau issues d'une indexation manuelle ou assistée (concepts) pour l'indexation et la recherche des documents vidéo.

Chapitre 2 : Formalisme de représentation des documents et base de connaissances

L'objectif de ce chapitre est d'introduire le formalisme des graphes conceptuels et les ontologies. Ces dernières sont utilisées pour présenter des descriptions partagées et plus ou moins formelles d'un domaine permettant ainsi d'ajouter une couche sémantique aux systèmes informatiques. L'indexation des documents par des mots, ou groupes de mots posent plusieurs problèmes.

Le formalisme de graphes conceptuels sera utilisé comme modèle opérationnel dans le cadre de notre proposition. Aussi, nous présentons les différents termes liés aux ontologies et les différents niveaux des ontologies.

Introduction	34
I. Les graphes conceptuels (GCs)	34
I.1. Généralités	34
I.2. Notations	35
I.3. Quelques définitions	35
I.4. Les graphes simples	37
I.5. Les graphes emboîtés typés	38
I.6. Les règles et contraintes	39
I.7. Synthèse	40
II. Les ontologies	40
II.1. Définitions	40
II.1.1. Vocabulaire contrôlé	40
II.1.2. Taxonomie	41
II.1.3. Thésaurus	41
II.1.4. Réseau sémantique	41
II.2. De la philosophie à l'informatique	42
II.2.1. Intelligence artificielle	42
II.2.2. Web sémantique	42
II.3. Les niveaux des ontologies	43
II.4. Composants des ontologies	44
II.5. Langages de spécification d'ontologie	44
II.5.1. Resource Description Framework : RDF	44
II.5.2. DAML+OIL	45
II.5.3. OWL (Ontology Web Language)	45
II.6. Ontologies et Recherche d'Information	45
Conclusion	48

Introduction

L'objectif de la modélisation des documents vidéo est de permettre une exploitation efficace de ceux-ci quelle que soit l'application envisagée. Si, par exemple, cette application est l'indexation et la recherche par le contenu, il est nécessaire que le formalisme de représentation doive permettre la définition d'une opération pour le calcul de la pertinence des réponses par rapport à un besoin d'information exprimé sous la forme d'une requête par un utilisateur.

Un formalisme de représentation met en place une représentation (linéaire ou sous forme des graphes) qui permet de décrire le document (structure et contenu). Un tel formalisme est facilement compréhensible par les êtres humains et avoir une capacité de représentation suffisante pour pouvoir être utilisé dans les diverses applications visées. L'indexation de documents, que nous proposerons, nécessite l'utilisation d'un vocabulaire contrôlé partagé par les acteurs du domaine. Ce vocabulaire se traduit par une ontologie évolutive en fonction du besoin de la communauté et de l'information existante.

Notre proposition utilisera donc le formalisme des graphes conceptuels ainsi que des ontologies pour la représentation des connaissances pertinentes dans les contextes visés. Ceux-ci sont présentés et étudiés dans les sections suivantes.

I. Les graphes conceptuels (GCs)

I.1. Généralités

Le formalisme des graphes conceptuels est un modèle de représentation de connaissances, [Sowa, 84]. C'est un ensemble de graphes bipartites, orientés et connexes. Les deux types de nœuds représentent le concept et la relation conceptuelle.

Les nœuds de type concept sont représentés par un couple comprenant un label de type et un référent. Par exemple : [Personne : Médecin] désigne un concept de type « personne » et identifié par son rôle « Médecin ». Les référents peuvent être des constantes correspondant à des identifiants uniques ou bien indéfinis désignés par (*) correspondant à un concept générique.

Les nœuds relations sont représentés par les types de relations utilisées. Par exemple, la relation spatiale «à gauche de » peut relier les deux concepts suivants :

$$[\text{Objet : *}] \rightarrow (\text{à gauche de}) \rightarrow [\text{personne : Médecin}]$$

Une relation est reliée au moins à un concept tandis qu'un concept peut être isolé. Un concept est constitué d'un type de concept (une idée ou représentation de l'esprit qui résume une multiplicité d'objets) et d'un référent. Ce dernier correspond à tout marqueur qui désigne un élément particulier dont le type est celui en question.

La **figure 1.4** montre un exemple de représentation sous forme graphique et sous forme textuelle avec le formalisme des graphes conceptuels.

Les concepts et les relations conceptuelles sont classés dans des treillis dits *treillis de concepts* et *treillis de relations*.

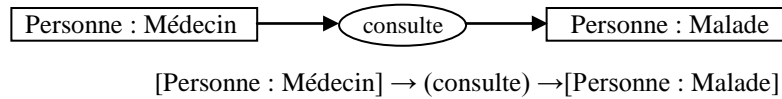


Figure 1. 4: Exemple de représentation avec les GCs



Figure 1. 5 : (a) Exemple de treillis de concepts, (b) exemple de treillis de relations

I.2. Notations

Nous rappelons dans cette section, quelques notations de base utilisées dans le formalisme des graphes conceptuels

Un graphe conceptuel g est défini par la donnée d'un triplet $(C(g), R(g), L(g))$, où :

- $C(g)$ est l'ensemble des concepts figurant dans le graphe g ;
- $R(g)$ est l'ensemble des relations reliant les concepts dans g ;
- $L(g)$ est un sous-ensemble de $C(g) \times C(g) \times R(g)$ qui donne pour chaque couple de concepts les relations qui les relient dans le graphe g .
- Un concept c est représenté par un couple, $c = [\text{type}(c) : \text{Réfèrent}(c)]$, où :
- $\text{Type}(c) = t$, est le type du concept, $t \in T$, T étant l'ensemble des types définis dans le contexte considéré ;
- $\text{Réfèrent}(c) = i$ est le réfèrent du concept c , $i \in I$, I étant l'ensemble des identifiants représentant les entités du monde réel que nous modélisons par des concepts.

Une relation est représentée par son type, $r \in R$ où R désigne l'ensemble des types de relations conceptuelles utilisées.

L'ensemble des concepts et des relations sont partiellement ordonnés par une relation de type spécifique / générique (\leq). Les types de concepts et les types de relations sont extraits d'une hiérarchie de types organisée en treillis (nommés respectivement treillis de concepts et treillis de relations). Ces treillis sont limités au sommet par le type universel \top et à la base par le type absurde \perp . Dans notre définition des graphes conceptuels, le treillis correspond à une structure arborescente.

Les graphes conceptuels peuvent aussi avoir d'autres graphes comme référents. Le graphe inclusif est alors défini comme étant le contexte du ou des graphe(s) inclus. Graphiquement, cela s'exprime en traçant un cadre autour des graphes inclus dans le contexte.

I.3. Quelques définitions

Dans [Mechkour, 95], un graphe conceptuel canonique et la représentation d'une base de graphes conceptuels nommée canon sont définis comme suit :

Définition 1 : un graphe conceptuel canonique est un graphe qui exprime une combinaison de concepts et de relation conceptuelle. La notion de graphe canonique a été introduite pour limiter la production des graphes conceptuels aux seuls graphes pouvant être interprétés comme ayant un sens dans le contexte de l'application.

Définition 2 : un canon est une représentation qui permet de définir une base de graphes conceptuels dans un domaine particulier. Il est défini par (T, I, C, B) , où :

- **T** est le treillis de types de concepts et de relations.
- **I** est l'ensemble des instances représentant les entités du monde réel que nous modélisons par les graphes conceptuels.
- **C** est la relation de conformité qui lie les éléments de **T** aux éléments de **I**
- **B** est la base canonique, dans laquelle tous les labels de types sont des éléments de **T** et tous les référents sont soit * (générique), soit des éléments de **I**.

Plusieurs règles fondamentales permettent de manipuler de manière cohérente les graphes conceptuels, [Sowa, 84] :

Règle 1 : Copie d'un graphe : Elle permet de copier un graphe dans un ou plusieurs autres graphes.

Règle 2 : Restriction des graphes : La restriction remplace dans un graphe canonique chacun des types par un de ses sous-types ou la valeur de son référent par un plus précis.

Règle 3 : Jointure de deux graphes : Nous appelons un concept commun à deux graphes un concept c figurant dans ces deux graphes. Étant donné un concept c commun à deux graphes a et b , un graphe d peut être construit en supprimant c dans a et en rattachant au concept c dans le graphe b toutes les relations « dépendantes » de a .

Règle 4 : Simplification des graphes : Lorsqu'une relation r relie deux concepts c_1 et c_2 plus d'une fois, nous disons qu'il y a des « relations redondantes ». Un graphe peut être simplifié en supprimant les relations redondantes, en laissant seulement une occurrence de la relation r entre c_1 et c_2 .

Règle 5 : Jointure Maximale : Un joint maximal de deux graphes est obtenu en joignant ces graphes par des sous-graphes compatibles maximaux des deux graphes. Les sous-graphes compatibles sont les plus grands sous-graphes des deux graphes qui ont une restriction en commun.

Règle 6 : Spécialisation : La restriction et la jointure sont des spécialisations parce qu'elles additionnent les informations des graphes initiaux. Si un graphe u peut être retiré d'un graphe v par une séquence de restriction et peut être joint avec d'autres graphes conceptuels, alors u est appelé une spécialisation de v .

Règle 7 : Généralisation : Si u est la spécialisation de v alors v est la généralisation de u . La généralisation est l'inverse de la spécification.

Règle 8 : Projection : Nous disons qu'il existe une projection d'un graphe b sur un graphe a s'il existe un sous graphe a' de a qui est une restriction de b . L'opération de projection est l'opération de base du modèle. Une projection d'un graphe G_1 dans un graphe G_2 peut être vue comme la recherche d'une fusion de l'information représentée par G_1 dans G_2 . On appelle alors le graphe G_2 comme étant une spécification de G_1 .

La **figure 1.6** présente un exemple de représentation avec le formalisme des graphes conceptuels.

Deux concepts faisant référence au même objet individuel ou au même ensemble sont dits co-référents. La dernière chose que nous dirons à propos des graphes conceptuels est qu'il en existe une forme linéaire qui permet de les mettre facilement en entrée d'un programme informatique. Cette notation linéaire nécessite l'introduction de variables pour mettre en évidence les coréférences.

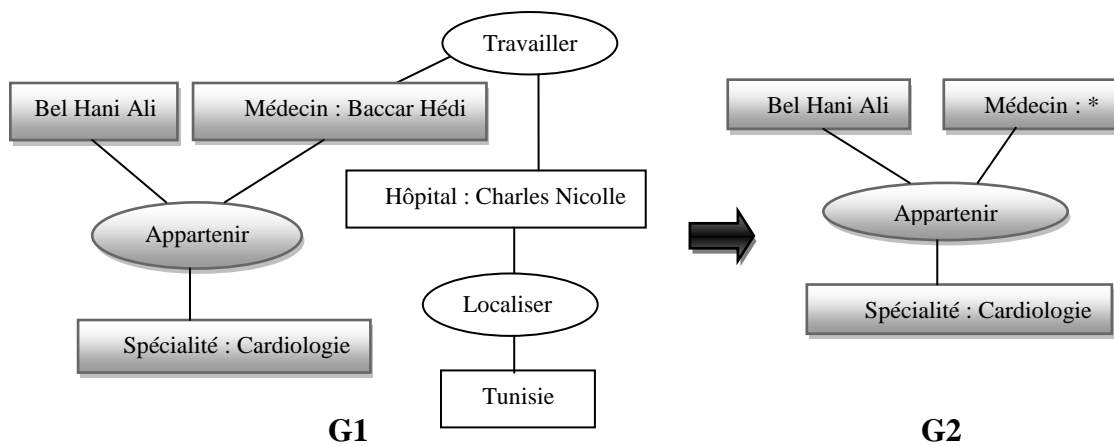


Figure 1. 6: Exemple d'opération de projection de GCs

I.4. Les graphes simples

Un graphe conceptuel simple (SG) est un graphe biparti étiqueté : l'une des classes de sommets, dite de sommets concepts, représente des entités, et l'autre, dite de sommets relations, représente des relations entre ces entités ou des propriétés de ces entités.

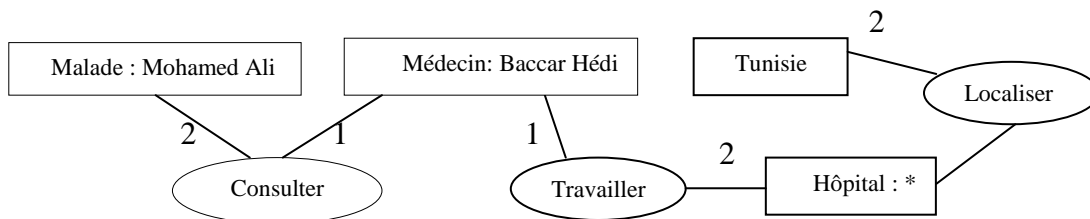


Figure 1. 7 : SG description du profil du médecin Baccar Hédi

Un sommet concept d'un SG est étiqueté par un couple $t : m$ où t est un type de concept (possiblement conjonctif) et m est un marqueur. Si le sommet représente une entité précise alors m est un identifiant représentant cette entité (le sommet est dit individuel), sinon m est noté $*$ (le sommet est dit générique). Le sommet [Médecin : Baccar Hédi] du graphe de la **figure 1.7** désigne Medecin, tandis que le sommet [Hôpital : *] désigne « un hôpital ». Un sommet relation est étiqueté

par une relation r et, si n est l'arité de r , n arêtes sont incidentes à ce sommet ; ces arêtes sont totalement ordonnées. Nous notons $G = (C_G, R_G, E_G, l_G)$ un SG où C_G est l'ensemble des sommets concepts, R_G l'ensemble des sommets relations, E_G l'ensemble des arêtes et l_G la fonction d'étiquetage des sommets et des arêtes.

Les étiquettes des sommets sont prises dans un vocabulaire appelé support . Un support (i.e. une ontologie simple) est défini comme une structure $O = (T_C, T_R, B, \sigma)$ où:

- T_C et T_R sont des ensembles ordonnés par spécialisation de types de concepts et de relations d'arité quelconque respectivement ($t' \leq t$ s'interprète par « t' est une spécialisation de t ») ;
- B est un ensemble de types bannis. Un type banni est un type conjonctif, par exemple {Humain, Objet} qui ne peut pas désigner une même entité (les sous-types d'un type banni sont eux-mêmes des types bannis).
- σ associe à chaque relation n -aire, sa signature, un n -uplet de types de concept qui définit le type maximal de chacun de ses arguments.

Un graphe G est cohérent par rapport à une ontologie $O = (T_C, T_R, B, \sigma)$ si:

- Les étiquettes des sommets relations et types des sommets concepts de G appartiennent respectivement à T_R et T_C ;
- Le type conjonctif de chaque individu – le type des sommets concepts génériques, et la conjonction des types des sommets concepts ayant un même identifiant pour les individuels – n'est pas sous-type d'un type banni dans B ;
- Les sommets relations respectent les signatures définies dans σ ;

La notion fondamentale pour comparer des SG est une application d'un SG dans un autre appelée projection. Intuitivement, l'existence d'une projection de G dans H montre que la connaissance représentée dans G est contenue dans H . Une projection π de G dans H est plus précisément une application de C_G dans C_H et de R_G dans R_H qui conserve les arêtes (si on a une arête entre r et c étiquetée i dans G alors on a une arête entre $\pi(r)$ et $\pi(c)$ étiquetée i dans H) et peut « spécialiser » les étiquettes des sommets. Le résultat fondamental sur lequel le formalisme SG s'appuie est l'équivalence entre l'existence d'une projection et la déduction logique sur les formules associées par une transformation ϕ en logique des prédicats aux SG.

I.5. Les graphes emboîtés typés

Le modèle des Nested Conceptual Graphs (NCG) typés (TNG) a été introduit pour contextualiser des connaissances, [Chein et al., 07].

Cette classe est récursivement définie par :

1. Un graphe emboîté typé est obtenu à partir d'un SG en ajoutant à chaque sommet concept c un troisième champ, égal à **.
2. Soit G un graphe emboîté typé, c_1, c_2, \dots, c_k des sommets concepts de G , n_1, n_2, \dots, n_k des types d'emboîtement, et G_1, G_2, \dots, G_k des graphes emboîtés. Le graphe obtenu en substituant l'ensemble $\{(n_1, G_1), (n_2, G_2), \dots, (n_k, G_k)\}$ au troisième argument ** de c_i est un graphe emboîté typé.

Nous ajoutons alors à la structure ontologique une arborescence de types d'emboîtements T_E . Les types d'emboîtement permettent de préciser la nature du lien qui lie un concept à son graphe emboîté (description, association, composition, historique...), l'ordre permet d'introduire une spécialisation entre eux (par exemple $Composition \leq Association$).

Graphiquement, les couples (n_i, G_i) associés au sommet c sont représentés par des boîtes à l'intérieur du sommet c , étiquetées par n_i et contenant le graphe G_i . L'exemple de la **figure 1.8** montre un sommet concept contenant deux emboîtements (de type Interactions et Relations).

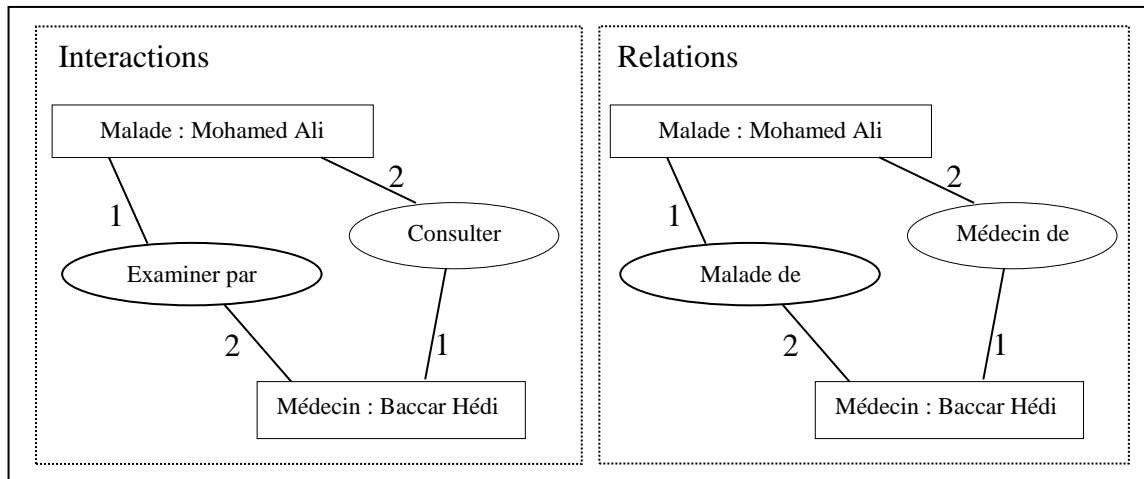


Figure 1. 8: Un graphe emboîté typé

I.6. Les règles et contraintes

Au formalisme « noyau » des graphes conceptuels simples, la famille SG ajoute deux types de connaissance : les règles et les contraintes, [Carloni, 08]. Une règle exprime une connaissance de la forme « si hypothèse alors conclusion », où hypothèse et conclusion sont deux SG. La **figure 1.9** présente un exemple de règles. Les pointillés lient certains sommets de l'hypothèse et de la conclusion. La règle exprime que si un médecin « consulte » un malade, alors ce malade « est examiné par » médecin. En appliquant la règle au SG de la **figure 1.9**, nous obtenons la relation supplémentaire « examiné par » entre [Médecin : Baccar Hédi] et [Malade : Mhamed Ali].

Les contraintes permettent d'étendre la notion de cohérence d'un graphe. Nous distinguons les contraintes positives des contraintes négatives :

- une contrainte positive a la même forme syntaxique qu'une règle mais exprime une obligation conditionnelle d'apparition d'un motif dans un graphe : « si une certaine connaissance est présente alors une autre doit l'être aussi ». Le graphe de la **figure 1.9** déclaré comme une contrainte imposerait qu'un graphe qui a la relation « consulter » entre un médecin et un malade doit également avoir la relation « examiner par » entre eux ;
- une contrainte négative a la même forme syntaxique qu'un SG, mais exprime qu'un certain motif « ne doit pas être trouvé ».

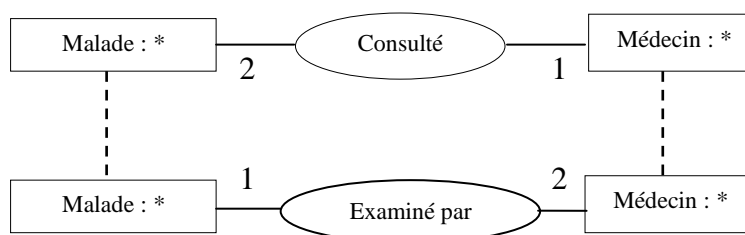


Figure 1. 9: Exemple de règle de SG

I.7. Synthèse

Le formalisme des graphes conceptuels est un formalisme de représentation de connaissance permettant la représentation de relation entre concepts. L'une des particularités de ce formalisme est de permettre la représentation de connaissance sous forme graphique.

Cet aspect graphique nous semble très important car les éléments (concepts et relations) du modèle peuvent être directement et clairement présentés à un utilisateur. Nous pensons, en effet, que des graphes conceptuels peuvent être facilement créés, compris et manipulés par des personnes n'appartenant pas au domaine informatique et qui ne connaissent pas l'aspect technique du modèle de représentation de connaissances.

Les relations entre les concepts veulent dire créer un lien sémantique entre eux (par exemple, un même personnage intervient dans deux scènes différentes). Avec cette définition, il est même possible d'étendre la notion de relation aux composants du document audiovisuel. En effet, la structure hiérarchique du document est en elle-même considérée comme une sorte de description relationnelle (appartenance, inclusion, composé de, etc....).

La modélisation conceptuelle permet de représenter les choses à l'utilisateur d'une manière différente. Ceci ne veut pas dire que le contenu ainsi décrit soit forcément différent [Charnois, 11] et [Raimbault *et al.*, 08].

II. Les ontologies

II.1. Définitions

II.1.1. Vocabulaire contrôlé

Un vocabulaire contrôlé, [The, 01], est une liste de termes définis généralement par une communauté afin de pouvoir décrire du contenu, et de rechercher l'information. Il est souvent utilisé pour les documents techniques ou plus généralement dans un domaine spécialisé. L'utilisation d'un vocabulaire contrôlé pour la description des documents facilite l'accès à l'information pour les utilisateurs par la rapidité et la délimitation d'un domaine de connaissances. Les résultats d'une recherche sont plus précis et pertinents par rapport à une description libre car un sujet sera décrit avec les mêmes termes.

Un vocabulaire contrôlé est utilisé dans les thésauri, les ontologies, les réseaux sémantiques, ainsi que toutes autres ressources décrivant un domaine avec des termes préférentiels. L'organisation du vocabulaire est traitée à différents niveaux, par exemple (relations entre les termes ou le multilinguisme).

II.1.2. Taxonomie

La structure la plus simple d'un vocabulaire contrôlé est la taxonomie. Il s'agit d'une hiérarchie de termes, organisée généralement avec la relation de spécialisation/généralisation. D'autres relations sont utilisées comme la composition mais dans une taxonomie, un seul type de relation est représenté [Texier, 05].

II.1.3. Thésaurus

Les termes d'un thésaurus servent à représenter ou à indexer des documents. Le thésaurus utilise un vocabulaire contrôlé, structuré et souvent restreint à un domaine particulier. En plus des relations de spécialisation (relation verticale) présentes dans une taxonomie, un thésaurus élargit le contexte d'un terme en ajoutant d'autres relations : terme interdit, terme préféré (relations horizontales). Cet élargissement sert au module de recherche de documents afin d'augmenter la pertinence avec des résultats correspondant à tout un contexte et non pas à un seul terme.

Un thésaurus est une sorte de dictionnaire hiérarchisé. Les termes normalisés sont reliés par des relations sémantiques. Il est organisé alphabétiquement, et propose généralement les définitions pour les termes, [Charlet *et al.*, 09].

Exemples de thésaurus

TT : POLLUTION DE L'EAU TG : POLLUTION
 TS : EUTROPHISATION
 TS : POLLUTION DES MERS
 TS : POLLUTION DE L'EAU SOUTERRAINE
 TA : EAU USEE
 TA : QUALITE DE L'EAU
 TA : LESSIVAT
 TA : LESSIVAGE DU SOL

Le thème TT : POLLUTION DE L'EAU pourrait ainsi être spécifié par au moins un terme descriptif choisi parmi les trois termes spécifiques (TS) ou parmi les termes génériques (TG), selon les besoins. Les termes associés (TA) seront par principe évités dans l'indexation, mais pourront être utilisés ultérieurement pour exploiter exclusivement tel ou tel type de contribution sans employer rigoureusement les termes propres de la description initiale.

II.1.4. Réseau sémantique

Le réseau sémantique est un outil qui simule la représentation humaine de la mémoire. C'est un modèle qui montre comment l'information pourrait être représentée en mémoire et comment nous pourrions accéder à ces informations, [Schwab, 05].

Un réseau sémantique est une représentation de l'ensemble des connaissances qu'un individu se construit pour un domaine spécifique. Les éléments de connaissance sont représentés par des nœuds et sont reliés par des liens associatifs ou sémantiques. Contrairement au thésaurus qui peut associer un terme à plusieurs contextes, dans un réseau sémantique un concept est défini de façon unique. Chaque concept est représenté par des attributs propres aux nœuds et par des relations qui l'associent aux autres nœuds.

Rastier, [Rastier, 04], rapproche les ontologies des réseaux sémantiques, et note que la principale différence réside dans la nouveauté de l'utilisation des ontologies.

Toutes ces définitions montrent la difficulté de distinguer tous ces concepts, ainsi que la subtile différence qui existe entre eux.

Définition : Une ontologie se rapproche d'un thésaurus par les relations qui la forment, ainsi que des réseaux sémantiques par le souhait de ne pas se limiter à des relations prédéfinies. Le principe ici est l'utilisation d'un vocabulaire contrôlé pour décrire un domaine traitant de ressources hétérogènes.

Avant de présenter les différents travaux sur la construction et l'enrichissement de l'ontologie, nous donnons quelques définitions du terme ontologie ainsi que son évolution dans le monde informatique.

II.2. De la philosophie à l'informatique

Il est difficile de donner une définition unique au terme ontologie car il apparaît dans plusieurs domaines. Le terme ontologie est apparu en premier en philosophie, il signifie « être » (du grec *ôn*, *onton*, participe présent du verbe *einai*) et discours, étude, sciences (*logos*). L'ontologie est la science ou la théorie de l'être, [Psyche *et al.*, 03].

Définition : « Science de l'être en tant qu'être, c'est-à-dire de l'être en général, de ses diverses espèces, de ses propriétés et de ses relations. Les modes de l'être sur lesquels elle discute sont : le possible, le réel et l'impossible, le potentiel et l'actuel, le contingent et le nécessaire, le déterminé et l'indéterminé, le fini et l'infini, le parfait et l'imparfait, la substance et le mode, l'essence et l'accident. », [Lame, 02].

Le monde de l'informatique s'est approprié plus tard ce terme.

II.2.1. Intelligence artificielle

McCarthy a été le premier dans le milieu de l'intelligence artificielle (IA), à s'intéresser aux ontologies de la philosophie, afin de construire des théories logiques des systèmes d'intelligence artificielle, en affirmant que pour pouvoir construire des systèmes intelligents fondés sur la logique, on devait construire une ontologie du monde afin d'énumérer tout ce qui existe, [Psyche *et al.*, 03].

Le terme ontologie a commencé à se répandre au début des années 90, une première définition est donnée par Neches et ses collègues, [Neches *et al.*, 91] :

Définition 1 : « Ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary. »

Gruber a ensuite donné sa définition, c'est celle qui est la plus citée en informatique, [Gruber, 93] :

Définition 2 : « An ontology is an explicit specification of a conceptualization. »

II.2.2. Web sémantique

Le Web sémantique est une extension du Web actuel où l'information est mieux définie dans le but de permettre aux machines et aux humains de coopérer. Dans le Web sémantique, une ontologie est vue comme un ensemble de connaissances, y compris le vocabulaire et les relations

sémantiques, avec quelques règles simples d'inférence et de logiques relatives à des sujets particuliers, [Ghaoui *et al.*, 05].

Une ontologie correspond à un vocabulaire contrôlé et organisé et à la formalisation explicite des relations créées entre les différents termes du vocabulaire. Cette formalisation peut être faite par des langages définis par la communauté du Web sémantique, [Simac, 11]. Une ontologie formelle peut être vue comme la distinction entre les entités du monde réel et les catégories utilisées pour décrire des entités (concepts, propriétés, relations). Une ontologie formelle, [Kassel, 09], est un développement systématique, formel et axiomatique de la logique de toutes les formes et modes d'existence. Les ontologies ont pour rôle principal la représentation d'un domaine.

Définition : «An ontology is a shared and common understanding of some domain that can be communicated between people and application systems», [Fensel, 00].

Elles sont utilisées dans divers domaines :

1. la description de l'information dans différentes spécialités (biologie, médecine...). Par exemple The **G**ene **O**ntology (GO), [Manser, 12] ;
2. la recherche d'information ;
3. le commerce électronique afin de permettre, par exemple, la communication entre les fournisseurs et les acheteurs ;
4. les environnements de formation à distance, [Merzougiet *al.*, 11].

II.3. Les niveaux des ontologies

Nous distinguons différents niveaux d'ontologies selon le but pour lequel elles sont conçues. Van Heijst distingue quatre types d'ontologies, [Zghal, 10] :

- les ontologies du domaine qui rassemblent les connaissances dans un domaine particulier ;
- les ontologies applicatives qui sont spécifiques et non réutilisables. Elles ont un domaine de validité restreint et correspondent à l'exécution d'une tâche ;
- les ontologies génériques qui expriment des conceptualisations valables dans différents domaines. Elles ne sont pas propres à un seul domaine ;
- les ontologies de représentations qui conceptualisent des primitives de langages de représentations des connaissances.

Guarino [Guarino, 99] donne une autre classification illustrée dans la **figure 1.10**.

Il définit les ontologies de haut niveau qui décrivent des concepts généraux comme le temps, une action, et indépendants d'un domaine particulier. Ce type ressemble aux ontologies génériques de Van Heijst. Les ontologies de domaine et les ontologies de tâche spécialisent les termes de l'ontologie de haut niveau. Enfin, les ontologies d'application décrivent des concepts qui dépendent de l'ontologie de domaine et de tâche.

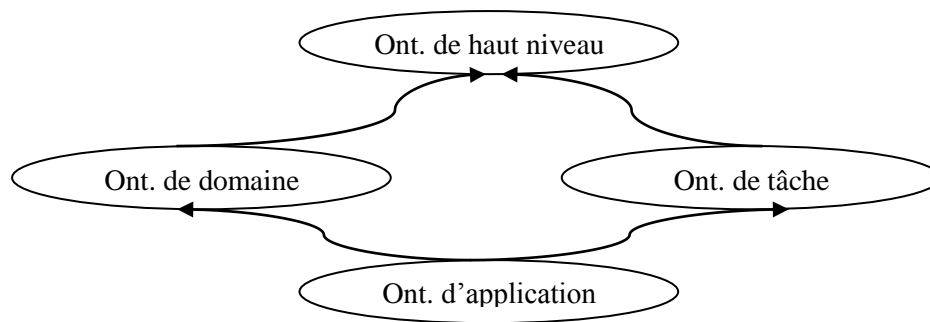


Figure 1. 10 : Les types d'ontologies

II.4. Composants des ontologies

D'après [Ahcene, 05], les connaissances traduites par une ontologie sont véhiculées à l'aide de cinq éléments : Concepts, Relations, Fonctions, Axiomes, Instances.

- *Les concepts* : sont appelés aussi termes ou classes de l'ontologie. Un concept est un constituant de la pensée (un principe, une idée, une notion abstraite) sémantiquement évaluable et communicable.
- *Les relations* : traduisent les associations qui existent entre les concepts présents dans une même partie analysée. Ces relations regroupent les associations suivantes : sous-classe-de (spécialisation, généralisation) ; partie-de (agrégation ou composition) ; associée-à ; instance-de ; est-un, etc.
- *Les fonctions* : sont des cas particuliers de relations dans lesquelles le nième élément de la relation est défini de manière unique à partir des n-1 premiers. Formellement, les fonctions sont définies ainsi : $F : C1 * C2 * \dots * Cn-1 \rightarrow Cn$. Comme exemple de relation binaire, nous avons la fonction mère-de.
- *Les axiomes* : sont utilisés pour structurer des phrases qui sont toujours vrais. Ils constituent des assertions, acceptées comme vraies, à propos des abstractions du domaine traduites par l'ontologie.
- *Les instances* : sont utilisées pour représenter des éléments.

II.5. Langages de spécification d'ontologie

II.5.1. Resource Description Framework : RDF

RDF est un langage d'assertion et d'annotations. Les assertions affirment l'existence de relations entre les objets. Elles sont donc adaptées à l'expression des annotations que l'on veut associer aux ressources du Web. Dans cette partie, nous parlerons principalement de RDF parce qu'il nous semble présenter des avantages déterminants pour la manipulation informatique, [Drame, 14].

RDF est un langage formel qui permet d'affirmer des relations entre des « ressources ». Il sera utilisé pour annoter des documents écrits dans des langages non structurés, ou comme une interface pour des documents écrits dans des langages ayant une sémantique équivalente (des bases de données, par exemple).

II.5.2. DAML+OIL

Ontology **I**nterchange **L**anguage(OIL) est une représentation basée sur le Web et la couche d'inférence pour des ontologies, qui combine les primitives de modélisation largement utilisés depuis le cadre des langages avec la sémantique formelle et des services fournis par le raisonnement des logiques de description.

DARPA **A**gent **M**arkup **L**anguage (DAML) Programme a débuté officiellement en août 2000. Le but de l'effort de DAML est de développer un langage et des outils pour faciliter le concept du Web sémantique. Pagels Michael est le DARPA gestionnaire de programme pour DAML.

DAML+OIL est la combinaison de langages OIL et DAML. Il hérite des avantages de ces deux langages. En conséquence, DAML+OIL est un langage très expressif et lisible par la machine ainsi que par un être humain avec une syntaxe basée sur RDF, [**Bechhofer et al., 2001**].

II.5.3. Ontology Web Language (OWL)

D'après la littérature, les langages mentionnés ci-dessus ont été utilisés pour développer des outils et des ontologies. Ils ne sont pas compatibles avec le Web Sémantique. D'où la nécessité d'un nouveau langage permettant de représenter les données sémantiques. Ce langage de représentation est l'OWL. Il est construit sur le modèle de données de RDF. Il nous donne les possibilités suivantes aux ontologies :

- capacité d'être distribué à travers plusieurs systèmes ;
- compatibilité avec des normes du Web pour l'accessibilité et l'internationalisation ;
- ouverture et extensibilité.

Il y a pas mal d'organismes utilisant OWL avec les nombreux outils disponibles. La majorité des systèmes qui ont utilisé DAML, OIL, ou DAML+OIL change maintenant à OWL.

II.6. Ontologies et Recherche d'Information

Les ontologies sont utilisées pour représenter des descriptions partagées et plus au moins formelles d'un domaine permettant ainsi d'ajouter une couche sémantique aux systèmes informatiques. C'est donc naturellement que des travaux sur l'intégration des ontologies dans les SRI se développent. De nos jours, de plus en plus d'ontologies sont conçues, accessibles et réutilisables. Intégrer les ontologies dans le module de recherche d'information consiste seulement à les choisir en considérant le domaine de connaissance qu'elles abordent, [**Baziz, 05**].

Une fois l'ontologie est choisie, la connaissance qu'elle représente peut être utilisée à différents niveaux dans le module de RI : au niveau de l'expansion des requêtes ou d'une façon plus poussée, au niveau de l'indexation.

Nous retrouvons dans [**Baziz, 05**] et [**Boughanem et al., 07**] deux principes de l'utilisation de l'ontologie dans le module de recherche d'information : soit dans l'expansion des requêtes soit dans le module d'indexation et de filtrage d'information comme schématisé dans la **figure 1.11**.

- Au niveau du *module d'indexation* des documents de la base : en s'associant à des techniques de **t**raitement **a**utomatique du **l**angage **n**aturel (TALN), les documents de la base seront résumés puis reliés à des concepts de l'ontologie. Si cette étape s'est déroulée convenablement (processus de désambiguïsation performant), la recherche serait plus aisée par la suite.

- Et, au niveau de la *reformulation de requêtes* utilisateurs, pour améliorer les requêtes utilisateurs.

Nous comprenons dès lors que l'amélioration des performances de ces moteurs de recherche en termes d'efficacité et de précision soit un souci pour les concepteurs des SRI. Les ontologies constituent pour cela une technologie jugée prometteuse. Nous allons voir que différentes utilisations des ontologies peuvent être envisagées.

L'indexation des documents par des mots, ou groupes de mots, pose plusieurs problèmes. Un des problèmes est qu'elle ne prend pas en compte les phénomènes de la langue que sont la synonymie et l'homonymie :

- L'indexeur privilégie en général un seul des mots, lorsque ces mots renvoient à un **même sens** mais différents, celui-là même qui apparaît dans le document. À supposer que l'utilisateur emploie dans sa requête un autre mot pour viser le même sens, il n'accède pas aux documents pertinents. La synonymie engendre du **silence** documentaire (le système a un mauvais rappel).
- Des mots identiques renvoient à des **sens différents**, l'effet inverse se produit. L'emploi du mot par l'utilisateur se traduit par une réponse contenant des documents traitant de l'ensemble des sens possibles... même s'il n'est intéressé que par un seul d'entre eux. L'homonymie engendre du **bruit** documentaire (le système a une mauvaise précision).

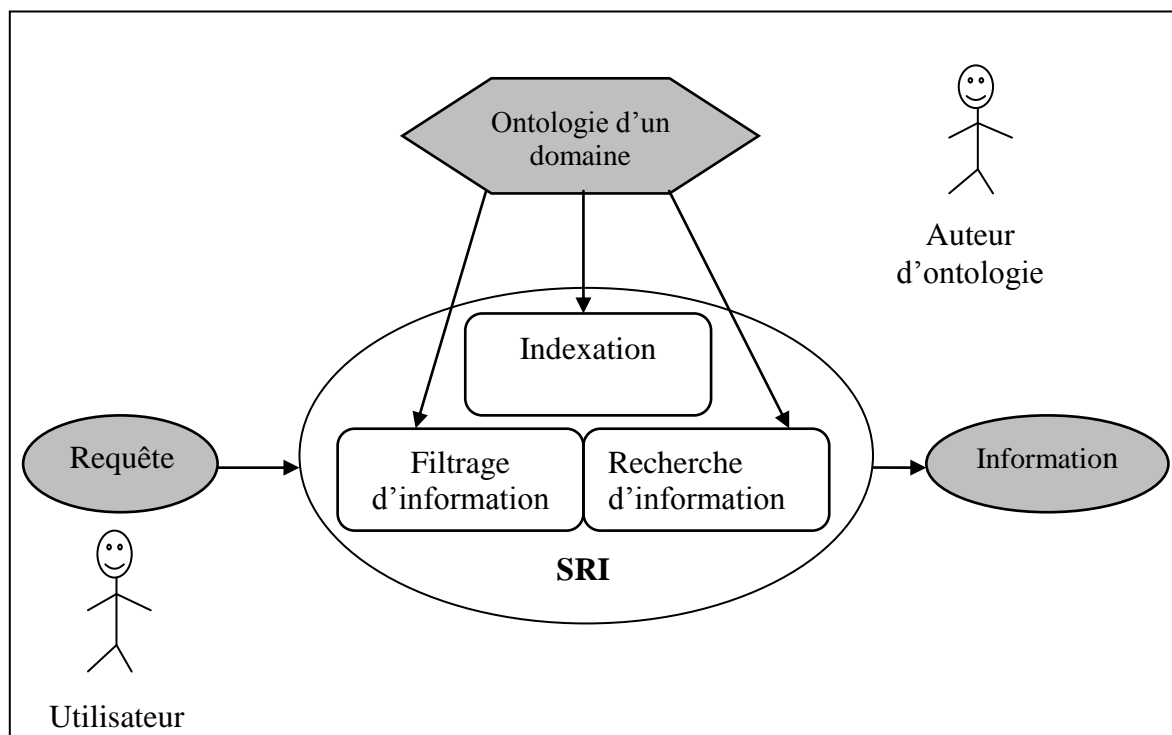


Figure 1. 11 : Principe d'intégration d'une ontologie au module de RI, [Said L'hadj, 09]

L'indexation des documents par des concepts est une solution pour pallier ces problèmes, tout en conservant les liens : concept/terme(s). Cette solution se situe dans l'axe des pratiques des documentalistes qui exploitent déjà des thésauri pour indexer des documents. L'ontologie (un

ensemble structuré de concepts) prend la place d'un thésaurus (un ensemble structuré de termes), [Benosman, 15].

L'apport d'une ontologie dans le cadre d'une indexation conceptuelle est double comme indiqué dans la **figure 1.12**.

- Elle permet à l'utilisateur de faire usage d'une autre terminologie que celle présente dans les documents.
- Elle permet au SRI d'aider l'utilisateur à reformuler ses requêtes sur la base d'une proximité sémantique.

Dans le cadre d'un module d'indexation et de recherche, l'enjeu d'utiliser une ontologie est de permettre à une large communauté d'utilisateurs de partager la signification de ses éléments, de façon à les utiliser correctement. Les utilisateurs de l'ontologie étant en premier lieu des êtres humains, et non des machines. Lors de module d'indexation et de recherche, l'ontologie est utilisée pour réaliser des inférences.

Dans ce cadre, Harrathi dans [Harrathi, 09] propose une approche statistique et ontologique d'indexation adaptée aux documents multilingues. Il propose une technique statistique exploitant les fréquences de mots afin d'extraire les termes des documents. Par la suite, il utilise le modèle d'ontologie afin d'associer les termes à leurs concepts (cf. **Figure 1.12**). Ce modèle est aussi exploité pour extraire les relations entre les concepts à partir des documents. Dans cette approche, Harrathi propose des algorithmes indépendants de la langue des textes pour reconnaître des concepts et des relations de l'ontologie dans les textes. Ces expérimentations ont résolu les problèmes rencontrés dans l'approche linguistique.

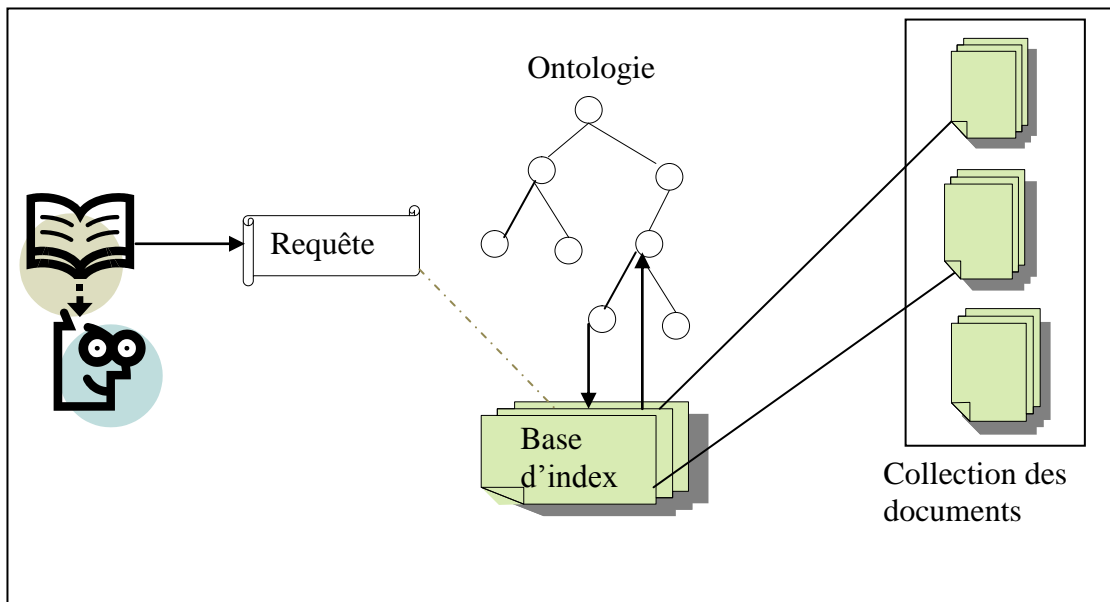


Figure 1.12 : Utilisation d'ontologie pour le module d'indexation

Conclusion

Nous avons détaillé dans les deux parties de ce chapitre l'utilité d'utiliser des ressources externes pour enrichir et élargir le vocabulaire de description du contenu vidéo. Les approches de modélisation décrites dans ce chapitre ne permettent de représenter et de décrire le contenu que partiellement. Ce manque d'expressivité de description se traduit par le manque d'exploitation des relations conceptuelles pour associer les différentes descriptions (concepts). Mise à part l'approche de modélisation en strates qui propose quelques relations simples et spécifiques qui permettent de regrouper les annotations sous formes graphiques, les approches existantes n'utilisent que de relations de type composition ou appartenance, liées à la structure du document.

L'indexation des DAV peut profiter des techniques permettant de créer et de manipuler des index en tant que véritables connaissances structurées. Pour cela, il faut des ontologies riches et capables d'assister efficacement le module de recherche.

Pour conclure, nous avons vu, dans ce chapitre, les documents audiovisuels et nous avons présentés le formalisme de graphe conceptuel. Aussi, nous avons déclaré la définition d'ontologie, ses composantes et les langages spécifiques.

Chapitre 3 : Recherche d'information

Afin d'améliorer le module de recherche, nous proposons dans une première étape de reformuler la requête utilisateur. Une deuxième étape consiste à calculer les mesures de correspondance entre la requête reformulée et les documents du corpus afin de bien satisfaire le besoin de l'utilisateur. Dans ce chapitre, nous présentons un panorama des travaux liés à ces deux étapes mentionnées.

Introduction	50
I.Reformulation des requêtes	50
I.1. Approche de reformulation	53
I.1.1. Reformulation automatique de requête	53
I.1.2. Reformulation interactive de requête	55
I.2. Utilité du profil utilisateur	55
I.2.1 Les techniques de construction des profils utilisateurs	56
I.2.2. Les techniques d'exploitation des profils utilisateurs	57
I.3. Expansion automatique des requêtes	57
I.4. Combinaison des présentations des requêtes	58
I.5. Réinjection de pertinence	58
II.Appariement document-requête	59
II.1. Modèles d'appariement	59
II.1.1. Le modèle booléen	59
II.1.2. Le modèle booléen pondéré.....	60
II.1.3. Le modèle vectoriel	60
II.1.4. Le modèle vectoriel étendu	61
II.1.5. Le modèle probabiliste	63
II.1.6. Modèle inférentiel	64
II.1.7. Le modèle différentiel bayésien	64
II.1.8. Les modèles de langage.....	65
II.2. Projection sémantique	66
II.3. Calcul de correspondance.....	66
II.4. Classement	66
II.5. Présentation	67
Conclusion	68

Introduction

La tâche principale d'un SRI est de sélectionner dans une collection de documents ceux qui sont susceptibles de répondre aux besoins d'information de l'utilisateur. Pour diminuer le "silence" (proportion des documents pertinents non retrouvés) et le "bruit" (proportion de documents non pertinents parmi ceux renvoyés), trois modules sont généralement mis en œuvre dans les SRI. Premièrement, un module d'analyse qui vise à prétraiter les documents du corpus. Deuxièmement, un module d'indexation visant à fournir une représentation compacte et sémantique des documents et des requêtes à l'aide des concepts et des relations sémantiques entre ces concepts. Et finalement, un module de recherche qui se compose de deux tâches essentielles : la reformulation des requêtes et l'appariement permettant de sélectionner les documents pertinents par rapport à une requête.

L'utilisateur, ayant un besoin en information, interroge le système de RI en lui soumettant une requête q , souvent sous la forme d'une liste de mots clés, et s'attend à recevoir comme réponse une liste des documents de la collection répondant potentiellement à cette requête, que nous appelons documents pertinents. Les SRI se distinguent entre eux par la façon d'interpréter et de représenter les documents et les requêtes.

Dans ce chapitre, nous présentons les différents travaux existants dans la littérature qui ont traité le module de recherche. Ce module est généralement composé de deux tâches. En première étape, il prend en charge la génération d'une nouvelle requête en utilisant des ressources externes telles que les ontologies et la requête initiale. Puis, il calcule la correspondance (appariement) entre la requête reformulée et les documents du corpus pour satisfaire le besoin de l'utilisateur.

I. Reformulation des requêtes

Le module de recherche est généralement présenté par la **figure 1.13**.

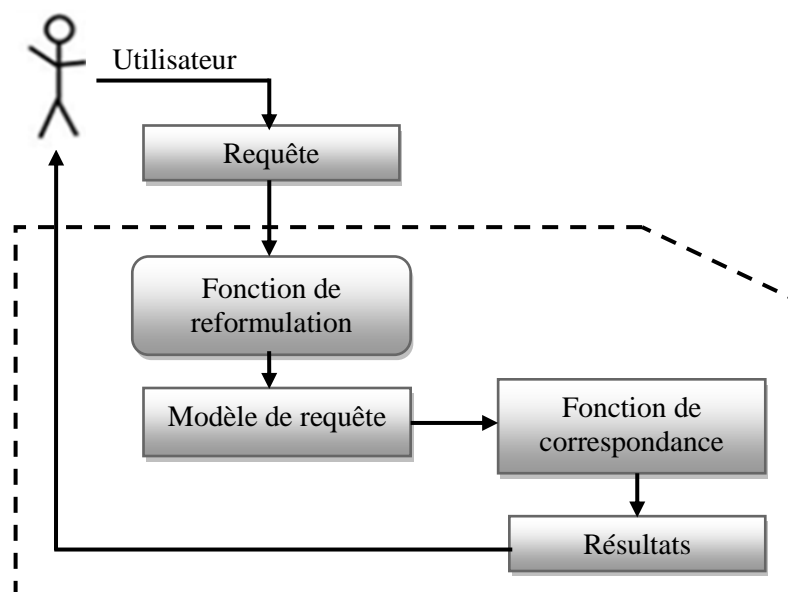


Figure 1.13 : Module de recherche

Parfois l'utilisateur est confronté à une situation difficile, il est incapable de trouver les mots précis pour exprimer son besoin en information. Alors, certains pour ne pas dire la majorité des documents qui lui sont retournés l'intéressent moins que d'autres.

Outre cela, certaines requêtes sont courtes, du coup, elles ne sont pas sémantiquement riches, pour que le SRI puisse retourner des documents pertinents. Afin de pallier ces problèmes, les chercheurs en RI se sont orientés vers l'intégration d'une étape supplémentaire dans le module de recherche : la reformulation ou expansion de requêtes. Elle consiste à modifier la requête initiale de l'utilisateur par l'ajout de termes significatifs et/ou la ré-estimation de leur poids, [Niang *et al.*, 13].

Si les termes rajoutés proviennent des documents de la collection, on parle de réinjection de pertinence (relevance feedback). En revanche, s'ils sont issus d'une ressource conceptuelle externe (ontologie, thésaurus ou dictionnaire), on parle, dans ce cas, de reformulation directe de requêtes, [Ben Aouicha, 09].

Un autre intérêt des ontologies est de permettre la désambiguïsation des termes de la requête. Plusieurs approches de désambiguïsation existent. Une approche consiste à choisir le concept dont les labels apparaissent le plus souvent dans les documents. Une autre consiste à réaliser un **profil utilisateur** (PU) et à choisir le concept le plus proche de son profil. Une troisième prend en compte le contexte de la recherche et les documents recherchés par l'utilisateur. Une quatrième améliore la désambiguïsation des sens des mots en utilisant la lemmatisation des mots au lieu d'utiliser la radicalisation. De plus, elle propose une méthode pour améliorer le rappel sans altérer la précision par l'utilisation des sous-concepts et super-concepts dans les différentes relations en respectant une certaine limite sur la profondeur des relations de subsomption, [Mataoui, 15].

Parmi les processus de reformulation existant dans la littérature, nous citons celui de Ralalason, [Ralalason, 10]. Ce processus peut être automatique ou manuelle suivant que le système ou l'utilisateur l'effectue. Ainsi, la reformulation de requêtes peut intervenir en particulier à deux niveaux :

- lors de la recherche initiale, si aucun document n'est retrouvé, le système réalise la reformulation ;
- si le lecteur n'est pas satisfait de la réponse du système après une première recherche, il peut reformuler sa requête.

Pour reformuler une requête, il propose quatre modes de reformulation : la généralisation, la spécialisation, la reformulation hybride qui combine sur une seule requête la généralisation et la spécialisation de la requête et enfin la prise en compte des concepts voisins, [Bouziri *et al.*, 16].

- **Généralisation de la requête** : dans le cas où le nombre de documents trouvés est insuffisant, le système ou l'utilisateur peut procéder à la généralisation de la requête en prenant comme nouveaux concepts annotant la requête le concept parent (plus général) de chaque concept de la requête.

- **Spécialisation de la requête** : dans le cas contraire, si le nombre de documents restitués est très important, le système ou l'utilisateur peut procéder à la spécialisation de la requête en prenant comme nouveaux concepts annotant la requête un ou plusieurs concepts spécifiques subsumés par chaque concept de la requête.
- **Reformulation hybride** (spécialisation et généralisation partielle) : la reformulation hybride est une façon de reformuler la requête en vue d'obtenir plus de précision sur les résultats. Cette reformulation se fait en spécialisant certains types de concepts et en en généralisant d'autres.
- **Prise en compte des concepts voisins** : la prise en compte des concepts voisins est une façon d'étendre le résultat d'une recherche.

En effet, les documents annotés par les concepts voisins d'un concept choisi par l'utilisateur peuvent l'intéresser. La différence par rapport la généralisation est qu'elle peut prendre en compte d'autres concepts.

[Dinh, 12] utilise, dans son travail de thèse, pour rechercher un document, deux étapes. Dans la première étape, il présente chaque requête sous forme d'un ensemble de concepts. Dans la deuxième étape, il calcule la pertinence du document par rapport à la requête.

- **Étape 1 : Représentation de la requête** : la requête Q peut être représentée formellement comme suit :

$$\begin{aligned} Q^s &= \{q_1^s, q_2^s, \dots, q_u^s\} \\ Q^w &= \{q_1^w, q_2^w, \dots, q_v^w\} \end{aligned} \quad (1.1)$$

où Q^s , Q^w sont respectivement l'ensemble de concepts et de mots simples, u et v sont respectivement le nombre de concepts et mots de Q , q_k^s est le k -ième concept et q_k^w est le k -ième mot de Q .

- **Étape 2 : Calcul de la pertinence** : la mesure de pertinence du document D_i vis-à-vis de la requête Q considère dans notre cas deux principaux facteurs : (1) l'adéquation du sens des concepts dans le document et de la requête, (2) la spécificité des concepts dans le document. Formellement, la similarité sémantique entre la requête Q et le document D_i , dénotée $RSV(Q, D_i)$, est donnée par :

$$RSV(Q, D_i) = RSV(Q^w, D_i^w) + RSV(Q^s, D_i^s) \quad (1.2)$$

où

- $RSV(Q^w, D_i^w)$ est la mesure TF-IDF classique basée sur les mots simples
- $RSV(Q^s, D_i^s)$ est la mesure de pertinence basée sur le sens du terme désignant un concept du document vis-à-vis de la requête, calculée comme suit :

$$RSV(Q^s, D_i^s) = \sum_{q_k^s \in Q^s} (1 + h(q_k^s)) * TF_i(q_k^s) * IDF(q_k^s) \quad (1.3)$$

où

- TF_i est la fréquence normalisée du mot q_k^w ou du concept q_k^s dans le document D_i ,
- IDF est la fréquence inverse de documents de q_k^w ou q_k^s dans la collection,
- $h(q_k^s)$ est la spécificité du terme q_k^s associée à son propre sens dans la requête, calculée comme suit :

$$h(q_k^s) = \frac{\text{niveau}(q_k^s)}{\text{MaxDepth}} \quad (1.4)$$

où le $\text{niveau}(q_k^s)$ correspond au niveau de la profondeur de q_k^s et MaxDepth est la profondeur maximale des concepts dans le thésaurus.

I.1. Approches de reformulation

Nous distinguons principalement trois approches de reformulation de requêtes : une approche basée sur un processus automatique, une approche basée sur un processus manuel et une autre basée sur un processus interactif. Nous allons détailler dans les paragraphes suivants ces trois approches et nous présentons les principaux travaux développés, [Niang *et al.*, 13] et [Audeh, 14].

I.1.1. Reformulation automatique de requête

La reformulation automatique de requête induit un processus d'expansion et/ou de repondération de la requête initiale en utilisant des critères de choix définis sans intervention de l'utilisateur. Ce type de reformulation peut être défini dans un contexte global, basé sur le thésaurus, ou alors local, basé sur les résultats de la recherche en cours.

I.1.1.1. Reformulation basée sur le contexte global

Cette stratégie de recherche fait référence à l'exploitation d'informations préalablement établies dans la collection, et non dépendantes de la recherche en cours, en vue de réaliser la reformulation. Ceci fait alors appel essentiellement à l'utilisation de thésaurus, [Benosman, 15].

- *Utilisation d'un thésaurus manuel*

Le principe fondamental est d'ajouter à la requête initiale, les termes voisins définis dans le thésaurus et sélectionnés par l'application d'un seuil et d'un algorithme de choix.

Généralement, la recherche d'information est effectuée selon les principales étapes suivantes :

- Expansion de requête en utilisant les liens sémantiques prédéfinis dans le thésaurus. En fait, on intègre à la requête utilisateur l'ensemble des termes qui traduisent la couverture sémantique de chacun de ses termes.
- Calcul de pertinence des documents selon un mécanisme d'activation propagation basé sur le modèle connexionniste.

- *Utilisation d'un thésaurus automatique basé sur la similarité*

L'expansion de requête basée sur un thésaurus est construite de façon automatique, modélisant des liens de similarité entre termes. À chaque terme, un descripteur vectoriel est associé. L'expansion de requête est alors effectuée selon les étapes suivantes :

- représenter sous forme vectorielle, la requête initiale Q ;

- utiliser le thesaurus pour calculer la similarité ;
- ajouter à la requête des termes.

- **Utilisation d'un thesaurus basé sur le contexte**

L'idée essentielle est d'étendre une requête par intégration de termes de même contexte que ceux qui la composent. Dans ce cadre, les approches diffèrent relativement au principe adopté pour la définition d'un contexte de mot.

L'expansion de requête est effectuée comme suit :

- Construction préalable du thesaurus de contexte de termes ;
- Calcul de similarités vecteur contexte – requête ;
- Ajout des termes dont la valeur de similarité avec la requête est supérieure à un seuil déterminé

- **Reformulation basée sur une combinaison de thesaurus**

Cette méthode d'expansion de requête utilise un thesaurus combiné de manière à conjuguer leurs caractéristiques.

Plus précisément, trois types de thesaurus sont utilisés dans l'approche :

- Thésaurus manuel : Un terme y est représenté selon différentes taxonomies. On y associe un graphe sémantique où les nœuds représentent les termes et liens des relations de synonymie entre termes inter-taxonomies et intra-taxonomies.
- Thésaurus basé sur la cooccurrence : On évalue la cooccurrence de pseudo-phrases de taille fixe T dans des blocs adjacents de documents.
- Thésaurus basé sur le contexte linguistique : les mots sont classés par contexte grammatical verbe, sujet, adjectif etc., puis on calcule la cooccurrence relative entre deux termes, dans chaque classe, selon une formule appropriée.

1.1.1.2. Reformulation basée sur le contexte local

Dans le cas de cette stratégie de recherche plus connue sous l'expression anglaise « *ad hoc feedback* », les informations utilisées pour la reformulation de requête dépendent en grande partie de la recherche en cours : documents retrouvés, termes et poids associés, [Dinh, 12].

À l'origine, les travaux relatifs à l'utilisation de cette stratégie consistent essentiellement en l'application de techniques de classification de termes issus des n tops documents retrouvés. Actuellement, de nouvelles techniques sont mises en œuvre en vue d'analyser le contexte local de la recherche et de l'exploiter pour l'expansion de requête.

1.1.1.3. La reformulation par injection de pertinence

La reformulation de requête par injection de pertinence est plus connue sous le nom de *Relevance Feedback*. Cette méthode permet une modification de la requête initiale, sur la base des jugements de pertinence de l'utilisateur sur les documents restitués par le système. La *relevance feedback* est une forme de recherche évolutive et interactive. Son principe fondamental est d'utiliser la requête initiale pour amorcer la recherche d'information puis

exploiter itérativement les jugements de pertinence de l'utilisateur afin d'« ajuster » la requête par expansion ou repondération.

La nouvelle requête obtenue à chaque itération de feedback, permet de « corriger » la direction de recherche dans le fond documentaire, et ce, dans le sens des documents pertinents. Un nombre considérable de travaux se sont intéressés à l'intégration de techniques de relevance feedback à des modèles de recherche de base, [Dinh, 12].

1.1.1.4. Reformulation dans le modèle vectoriel

Les stratégies de reformulation développées dans le modèle vectoriel induisent une repondération de requête avec expansion. La reformulation consiste alors à orienter le vecteur requête vers les vecteurs documents pertinents et de l'éloigner des vecteurs documents non pertinents.

1.1.1.5. Reformulation dans le modèle connexionniste

La relevance feedback a été également expérimentée pour améliorer les résultats de la recherche d'information dans un modèle connexionniste, [Zemirli, 08].

Ce modèle est basé sur un réseau à deux couches : couche de mots et couche de documents. Les connexions entre ces deux couches sont bidirectionnelles et pondérées par des formules classiques.

L'interrogation du réseau à l'aide d'une requête, provoque la propagation des signaux d'activation depuis les neurones termes de la requête jusqu'aux neurones documents. Le jugement de pertinence est basé sur une estimation interne ; les documents dont le niveau d'activation est supérieur à un seuil sont considérés pertinents. L'itération feedback correspond alors à la poursuite de la recherche par propagation des signaux d'activation à travers les connexions inverses, depuis les neurones documents pertinents jusqu'aux neurones termes. Il en résulte le calcul d'une nouvelle activation des neurones termes qui correspond ainsi à la constitution d'une nouvelle requête.

1.1.2. Reformulation interactive de requête

À la différence de la reformulation automatique, l'approche interactive (ou par réinjection de pertinence et/ou non-pertinence) exploite uniquement un sous-ensemble de documents sélectionnés parmi les premiers résultats obtenus de l'exécution de la requête initiale. Son principe fondamental est d'utiliser cette requête pour amorcer la recherche, puis modifier celle-ci à partir des jugements de pertinence et/ou de non pertinence de l'utilisateur, soit pour re-pondérer les termes de la requête initiale, [Tamine *et al.*, 03], soit pour y ajouter (resp. supprimer) d'autres termes contenus dans les documents pertinents (resp. non pertinents). La nouvelle requête ainsi obtenue à chaque itération de feedback, permet de corriger la direction de la recherche dans le sens des documents pertinents.

1.2. Utilité du profil utilisateur

Une même requête émise par deux utilisateurs dont les besoins et les intentions sont différents, produit les mêmes résultats. Pour répondre aux problèmes de la surcharge informationnelle et pour pouvoir discriminer les utilisateurs en fonction de leurs besoins

spécifiques, de plus en plus de systèmes proposent des techniques de personnalisation. Parmi ces systèmes, nous citons ceux qui permettent à l'utilisateur de personnaliser les rubriques et l'apparence de ses pages d'accueil (MyYahoo², iGoogle³), ceux qui font des recommandations aux utilisateurs en fonction de l'historique de leurs actions (Amazon⁴, dailymotion⁵, etc.) ou encore les applications de vidéo à la demande, [Kim *et al.*, 05] et d'IPTV, [Al-Hezmi *et al.*, 06].

La personnalisation des données a été abordée dans différents domaines comme la RI, les **B**ases de **D**onnées (BD) ou l'**I**nteraction **H**omme-**M**achine (IHM). Dans le domaine de la RI, l'utilisateur fait partie du processus de personnalisation qui est vu comme un processus incrémental et interactif dans lequel l'utilisateur décide à chaque pas quels sont les éléments qu'il aime et quels sont ceux qu'il n'apprécie pas, [Kostadinov, 08], [Leung *et al.*, 12] et [Tchuente, 13]. La personnalisation est ainsi définie comme un apprentissage réalisé à partir des préférences rendues par les utilisateurs à l'issue de la présentation des résultats successifs du système.

La requête contient en général l'ensemble des critères considérés nécessaires à produire des données pertinentes. La personnalisation de l'information se fait par l'extension des langages de requêtage (généralement SQL) ou par un processus de reformulation de requêtes qui utilise les PUs, [Radinsky *et al.*, 13].

La personnalisation de l'information se fait généralement par la prise en compte d'un ensemble de connaissances, désignées souvent par les notions de *profil* et de *contexte*, qui décrivent les utilisateurs, leurs préférences et l'environnement dans lequel ils se trouvent.

Le PU est une mise en facteur des préférences de l'utilisateur. Pour la majorité des approches, le profil contient la description du centre d'intérêt de l'utilisateur. Parmi les formalismes utilisés pour le représenter, nous citons les vecteurs de mots clés (éventuellement pondérés), les ontologies, les fonctions d'utilité, les ensembles de prédicats pondérés ou encore les expressions logiques à la clause de Horn, [Latour, 14].

De même, le PU peut également contenir des données personnelles, des préférences sur la qualité des données et sur la sécurité. Dans ce cadre, la classification, l'organisation et la structuration des données du profil est un élément clé de la personnalisation.

I.2.1 Les techniques de construction des profils utilisateurs

Un problème important qui se pose dans le cadre des applications personnalisées est la construction des profils utilisateurs. La grande diversité des approches existantes nécessite l'identification d'un ensemble de paramètres qui permettant de les classer. Parmi les caractéristiques d'un processus de construction des profils, nous avons trouvé :

²MyYahoo : <http://cm.my.yahoo.com/>

³iGoogle : <http://www.google.fr/ig>

⁴Amazon : <http://www.amazon.com/>

⁵Dailymotion : <http://www.dailymotion.com>

- **Le moment où le profil est construit** : Généralement le processus de construction de profils utilise des techniques assez coûteuses basées par exemple sur l'apprentissage automatique ou sur la théorie des graphes. Par conséquent, la construction des profils est souvent faite en dehors des sessions d'interaction entre l'utilisateur et le système, [Kostadinov, 08].
- **L'implication de l'utilisateur** : Selon ce critère, les approches se divisent en deux groupes : celle qui n'impliquent pas l'utilisateur dans la construction du profil et celles qui impliquent l'utilisateur en lui demandant de remplir un formulaire (Yahoo) ou de corriger manuellement ses préférences.
- **Le type de sources de données utilisées pour la construction des profils** : Différents types de données peuvent être exploités pour la construction des PUs : données d'usage, données de contenu, données décrivant la structure et l'organisation des documents ou encore les données provenant de l'utilisateur.
- **La manière de mise à jour des profils** : Lorsque de nouvelles données sont collectées sur l'utilisateur ou sur son comportement, le profil doit être mis à jour pour tenir compte de ces données. Nous distinguons, selon ce critère, des systèmes qui mettent à jour uniquement les parties du PU concernées par les nouvelles données et les approches qui construisent à nouveau le PU.

I.2.2. Les techniques d'exploitation des profils utilisateurs

Une fois le profil utilisateur construit, il est utilisé par les systèmes d'information afin de fournir à l'utilisateur des données pertinentes, dans le format le plus approprié, au bon moment et au bon endroit. Pour répondre aux besoins de la personnalisation, différentes approches ont été développées. Nous nous focaliserons ici sur les techniques de personnalisation utilisées dans le domaine des Bases de Données. Les systèmes de bases de données proposent des langages de requêtes avancés qui supportent la notion de préférence. Il en est ainsi pour les langages comme Preference SQL ou SQL /f ou pour les opérateurs de type « Best of », et « Skyline ».

La reformulation de la requête peut utiliser de nouveaux critères de sélection contenus dans le profil utilisateur. Le PU est composé d'un ensemble de prédicats pondérés. Le poids d'un prédicat exprime son intérêt relatif pour l'utilisateur. Il est spécifié par un nombre réel compris entre 0 et 1. La reformulation d'une requête à l'aide d'un profil consiste à ajouter dans la requête initiale les prédicats de sélection les plus pertinents, [Saulier *et al.*, 14].

I.3. Expansion automatique des requêtes

L'expansion directe de la requête consiste à rajouter à la requête initiale des termes issus de ressources linguistiques existantes ou bien de ressources construites à partir des collections, [Plu *et al.*, 11]. Plus précisément, le but est d'utiliser un vocabulaire contrôlé issu de ressources externes. Nous pouvons alors utiliser des ontologies linguistiques. Nous pouvons également ajouter à la requête des variantes morphologiques des termes employés par l'utilisateur. Le but de ce mécanisme est d'assurer la restitution des documents indexés par des variantes des termes composant la requête, [Niang *et al.*, 13].

Les associations établies manuellement traduisent généralement des relations de synonymie et de hiérarchie. Les thésauri construits manuellement sont un moyen efficace pour l'expansion de requête. Cependant, leur construction et la maintenance des informations sémantiques qu'ils contiennent sont coûteuses en temps et nécessitent le recours à des experts des domaines considérés. Pour cette raison, ils restent peu utilisés par les SRI, [Audeh, 14].

En ce qui concerne la seconde catégorie de ressources, elles sont construites en s'appuyant sur une analyse statistique des collections. Il s'agit de chercher des associations de termes afin d'ajouter des termes voisins à la requête. Il existe aussi d'autres méthodes entièrement automatiques telles que le calcul des liens contextuels entre termes, [Claveau *et al.*, 04] et la classification automatique de documents.

Les associations créées automatiquement sont généralement basées sur la cooccurrence des termes dans les documents. Les liens inter-termes renforcent la notion de pertinence des documents par rapport aux requêtes.

I.4. Combinaison des présentations des requêtes

Plusieurs approches de RI utilisent une seule représentation de requête comparée à plusieurs représentations de document (algorithmes multiples de recherche). Il a été montré dans [Lee *et al.*, 99] qu'une recherche plus efficace peut être atteinte en exploitant des représentations multiples de requêtes ou des algorithmes de recherche différents ou encore en utilisant différentes techniques de réinjection.

Une combinaison des représentations de requêtes peut augmenter le rappel d'une requête, tandis que la combinaison des algorithmes de recherche peut augmenter la précision. La base théorique de la combinaison des évidences a été présentée par Ingwersen, [Ingwersen, 94]. Il a en particulier montré que des représentations multiples d'un même objet, par exemple une requête, permettent une meilleure perception de l'objet qu'une seule bonne représentation.

Cependant, il est important que chacune des sources d'évidences utilisées fournisse non seulement un point de vue différent sur l'objet, mais que ces points de vue aient différentes bases cognitives. Les représentations multiples d'une requête peuvent donner différentes interprétations du besoin en information.

Des approches de combinaison de multiples représentations de requêtes sont proposées dans [Tamine, 00] et [Dinh, 12]. Elles sont basées sur les algorithmes génétiques, plus précisément, elles proposent d'utiliser une population de requêtes qui évolue à chaque étape de la recherche et tente de récupérer le maximum de documents pertinents.

I.5. Réinjection de pertinence

Le processus de réinjection de pertinence, comporte principalement trois étapes : l'échantillonnage, l'extraction des évidences et la réécriture de la requête, [Bouramoul, 11].

- L'échantillonnage : cette étape permet de construire un échantillon de documents à partir des éléments jugés par l'utilisateur. Cet échantillon est caractérisé par le nombre d'éléments jugés et le nombre d'éléments jugés pertinents.

- L'extraction des évidences est l'étape la plus importante, elle consiste en général à extraire les termes pertinents qui serviront à l'enrichissement de la requête initiale.
- La réécriture de la requête consiste à construire une nouvelle requête en combinant la requête initiale avec les informations extraites dans l'étape précédente.

Le processus général de la réinjection de pertinence peut être renouvelé plusieurs fois pour une même séance de recherche : on parle alors de la réinjection de pertinence à itérations multiples. D'une manière générale, la phase d'échantillonnage ne présente pas de problématique spécifique. Le seul point abordé à ce niveau concerne le nombre d'éléments à évaluer pour pouvoir effectivement constituer un échantillon représentatif.

La problématique principale de la réinjection de pertinence réside dans les deux autres phases : l'extraction des termes (ils sont alors pondérés pour sélectionner les éléments les plus pertinents) et la réécriture de la requête avec repondération des termes.

Dans la plupart des approches de la littérature, les deux phases sont effectuées avec des méthodes de pondération des termes similaires. Cependant certaines méthodes et particulièrement celles basées sur le modèle probabiliste, utilisent des méthodes de pondération différentes.

II. Appariement document-requête

La comparaison entre le document et la requête permet de calculer une mesure appelée **pertinence système**, supposée représenter la pertinence du document vis-à-vis de la requête. Cette valeur est calculée à partir d'une fonction de similarité notée $RSV(Q,D)$ (**R**etrieval **S**tatus **V**alue), où Q est une requête et D est un document, [Ben Aouicha, 09]. Cette mesure tient compte du poids des termes dans les documents. D'une façon générale, l'appariement document-requête et le modèle d'indexation permettent de caractériser et d'identifier un modèle de recherche d'information. L'ordre dans lequel les documents susceptibles de répondre à la requête sont retournés est important. En effet, l'utilisateur se contente généralement d'examiner les premiers documents renvoyés (les 10 ou 20 premiers). Si les documents recherchés ne sont pas présents dans cette tranche, l'utilisateur considérera le SRI comme mauvais vis-à-vis de sa requête.

Dans ce qui suit, nous présentons les principaux modèles, qui ont par la suite été repris dans le cadre de la recherche d'information sémantique.

II.1. Modèles d'appariement

II.1.1. Le modèle booléen

Le modèle booléen, [Salton, 71], est historiquement le premier modèle de RI, et est basé sur la théorie des ensembles. Un document est représenté par une liste de termes (termes d'indexation). Une requête est représentée sous forme d'une équation logique. Les termes d'indexation sont reliés par des connecteurs logiques ET, OU et NON.

Le module de recherche mis en œuvre consiste à effectuer des opérations sur l'ensemble de documents afin de réaliser un appariement exact avec l'équation de la requête. L'appariement exact est basé sur la présence ou l'absence des termes de la requête dans les documents.

La décision binaire sur laquelle est basée la sélection d'un document ne permet pas d'ordonner les documents renvoyés à l'utilisateur selon un degré de pertinence, [Pal *et al.*, 11], [Drame, 14] et [Abbes, 15].

II.1.2. Le modèle booléen pondéré

Le modèle booléen a été étendu avec un nouvel opérateur binaire non commutatif, appelé "contains", [Tortman *et al.*, 03]. Le premier opérande est de type XPath et le second est une expression booléenne. Ce modèle permet aux requêtes d'être complètement spécifiées en termes de contenu et d'information structurée, basée sur XPath. La recherche consiste à extraire le titre et le convertir en requête booléenne, les éléments considérés comme pertinents sont par la suite classés selon la somme OkapiBM25.

Dans [Boubekeur *et al.*, 07], les auteurs utilisent dans une combinaison de méthodes probabilistes utilisant une régression logistique avec une approche basée sur le modèle booléen, pour évaluer la pertinence des documents et des éléments. La valeur de probabilité de pertinence R d'un composant C (élément) est calculée comme étant le produit des probabilités de la pertinence de C vis-à-vis la requête Q_{bool} présentée par un modèle booléen et de la pertinence de C vis-à-vis la requête Q_{prob} présentée par un modèle probabiliste. La formule est présentée ci-dessous :

$$p(R|Q, C) = P(R|Q_{bool}, C) P(R|Q_{prob}, C) \quad (1.5)$$

Cette combinaison permet de restreindre l'ensemble des documents pertinents aux documents ayant une valeur booléenne égale à 1 tout en leur attribuant un rang basé sur un calcul probabiliste. Ces deux types d'extension permettent de surmonter les limites des modèles booléens au niveau du tri des résultats.

II.1.3. Le modèle vectoriel

C'est un modèle qui préconise la représentation des requêtes utilisateurs et des documents sous forme de vecteurs, dans l'espace engendré par tous les termes d'indexation, [Salton, 89]. D'une manière formelle, les documents (D_j) et les requêtes Q sont des vecteurs dans un espace vectoriel des termes d'indexation (t_1, t_2, \dots, t_T) de dimension T et représentés comme suit :

$$D_j = [d_{j1}, d_{j2}, \dots, d_{jT}], Q = [q_1, q_2, \dots, q_T] \quad (1.6)$$

Où d_{ji} et q_i sont respectivement les poids des termes t_i dans le document D_j et la requête Q . D'après ce modèle, le degré de pertinence d'un document relativement μ a une requête est perçu comme le degré de corrélation entre les vecteurs associés. Ceci nécessite alors la spécification d'une fonction de calcul de similarité entre vecteurs mais également d'une fonction de pondération des termes. La plus répandue est celle de Sparck et Needham, [Sparck-Jones *et al.*, 72], qui définit le poids d'un terme t_i dans un document d_j comme suit :

$$d_{ij} = tf_{ji} * idf_i \quad (1.7)$$

Où :

tf_{ji} : est la fréquence relative du terme t_i dans le document D_j .

idf_i : est l'inverse de la fréquence absolue du terme t_i dans la collection.

$idf_i = \log \frac{n}{n_i}$ avec n_i le nombre de documents contenant le terme t_i

Et N est le nombre total de documents dans la collection.

La fonction de similarité permet de mesurer la ressemblance des documents et de la requête. La mesure la plus répandue est celle du cosinus, [Salton, 71] :

$$RSV(Q, D_j) = \frac{\sum_{i=1}^T q_i d_{ji}}{\sqrt{\sum_{i=1}^T d_{ji}^2} \sqrt{\sum_{i=1}^T q_i^2}} \quad (1.8)$$

Le modèle vectoriel suppose l'indépendance entre termes. En effet, la représentation vectorielle considère chaque terme séparément alors qu'on peut avoir des termes qui sont en relation sémantique entre eux, [Martinet, 04] et [Abbes, 15].

II.1.4. Le modèle vectoriel étendu

Le modèle vectoriel étendu est une extension du modèle vectoriel proposé en RI traditionnelle, dans lequel l'information structurelle est séparée du contenu, [Mass et al., 04].

Mass et al. [Mass et al., 04] ont proposé un système de recherche où le score d'un terme t_i dans un document D , $w_D(t_i)$, est exprimé par le produit $tf * idf$ et la pertinence d'un document pour une requête donnée est calculée selon l'équation suivante :

$$p(Q, D) = \frac{\sum_{t_i \in Q \cup D} W_Q(t_i) * W_D(t_i)}{\|Q\| * \|D\|} \quad (1.9)$$

Avec $\|Q\|$ et $\|D\|$ sont respectivement les normes des vecteurs de la requête Q et du document D . Pour chaque terme de la requête (t_i , c_i), (t_i est le terme, c_i est le contexte du terme (i.e. le chemin de l'élément où apparaît le terme t_i)) on calcule son poids dans la requête $w_Q(t_i; c_i)$, son poids dans un contexte similaire dans le document $w_D(t_i, c_k)$, ainsi que la similarité entre les contextes $cr(c_i, c_k)$ où

$$cr(c_i, c_k) = \frac{1 + |c_i|}{1 + |c_k|} \quad (1.10)$$

avec $|c_i|$ est le nombre de balises dans un contexte donné de la requête et $|c_k|$ est le nombre des tags dans un contexte donné d'un document.

La formule précédente est par la suite étendue pour mesurer les similarités entre les fragments XML et le document. La formule mesurant la pertinence d'un document est la suivante :

$$p(Q, D) = \frac{\sum_{(t_i, c_i) \in Q} \sum_{(t_i, c_k) \in D} W_Q(t_i, c_i) * W_D(t_i, c_k) * cr(c_i, c_k)}{\|Q\| * \|D\|} \quad (1.11)$$

Les auteurs ont considéré par la suite une autre méthode de mesure de similarité entre document et requête en considérant que les différents contextes c_k sont d'égale similarité avec le contexte de la requête. La formule appliquée est alors la suivante :

$$p(Q, D) = \frac{\sum_{(t_i, c_i) \in Q} W_Q(t_i) * W_D(t_i) * W(c_i)}{\|Q\| * \|D\|} \quad (1.12)$$

où $w(c_i) = 1 + |c_i|$ est le poids du contexte c_i .

Mass *et al.* [Mass *et al.*, 04] proposent de distinguer les composants d'un document, ce qui conduit à considérer six différents index (*article*, *sec*, *ss1*, *ss2*, *p*, *ip1* et *abs*) et à appliquer sur chacun le module de recherche tout en adaptant le *tf*, *idf* au niveau des composants. Pour éviter la redondance des résultats, les auteurs ont proposé un algorithme de classification basé sur la comparaison des scores des nœuds avec ceux de leurs descendants et selon un seuil déjà fixe, on décide d'éliminer l'un des groupes.

De plus, pour surmonter le problème des index de différentes caractéristiques, les auteurs ont proposé la notion de pivot pour avoir un nouveau score au niveau de chaque composant C calculé comme suit :

$$score(Q, C) = DocPivot * S_a + (1 - DocPivot) * S_c \quad (1.13)$$

avec *DocPivot* une constante entre 0 et 1, S_a le score du document auquel appartient le composant et S_c le score propre du Composant. L'algorithme de recherche est alors le suivant :

1. Le système effectue des recherches indépendantes au niveau de chaque sous index_i. Le résultat de chaque recherche est l'ensemble R_i .
2. Tous les résultats issus des différentes recherches sont normalisés dans [0,1].
3. Le score final de différents résultats est calculé en fonction du score de l'article. La liste finale des résultats est la combinaison de tous les résultats R_i .

Dans [Geva, 04], Geva a proposé un modèle simple donnant de très bons résultats pendant les campagnes d'évaluation INEX 2003 et INEX 2004. Ce modèle est basé sur un fichier inverse pour l'indexation d'un document XML. La recherche est réalisée par propagation des scores des éléments feuilles. Ces derniers sont calculés comme suit :

$$L = N^{n-1} \sum_{i=1}^n \frac{t_i}{f_i} \quad (1.14)$$

Avec

- N un entier faible = 5 ;
- n : nombre de termes de la requête qui existent dans le nœud n.
- t_i : fréquence du i^{eme} terme de la requête dans le nœud n.
- f_i : fréquence du i^{eme} terme de la requête dans la collection.
- N^{n-1} : augmente le score des éléments ayant plusieurs termes de la requête.

La pertinence R des nœuds internes est calculée par la somme des différents descendants :

$$R = D(n) \sum_{i=1}^n L_i \quad (1.15)$$

avec L_i est le score du i^{eme} élément retourné et $D(n)=0.49$ si $n=1$ et 0.99 sinon.

Enfin, on trouvera d'autres approches utilisant le modèle vectoriel étendu dans [Kakade *et al.*, 05] et [Hubert, 05].

II.1.5. Le modèle probabiliste

Pour étendre le modèle probabiliste inférentiel aux documents XML, les probabilités doivent tenir compte de l'information structurelle. Une approche est d'utiliser des probabilités conditionnelles sur les chemins des documents, avec par exemple $P(d|t)$ devenant $P(d|p$ contains $t)$, où d représente un document ou une partie de document, t est un terme et p est un chemin dans l'arbre structurel de d .

Une méthode d'augmentation basée sur le modèle probabiliste est proposée par Fuhr *et al.* dans [Fuhr *et al.*, 03]. Cette méthode est basée sur le langage de requêtes XIRQL, et a été implémentée au sein du moteur de recherche HyRex.

Dans cette approche, les nœuds sont considérés comme des unités disjointes. Tous les nœuds feuilles ne sont cependant pas indexés (car d'une granularité trop fine). Dans ce cas-là, les termes sont propagés jusqu'au nœud indexable le plus proche. Afin de préserver des unités disjointes, on ne peut associer à un nœud que des termes non reliés à ses nœuds descendants.

Le poids de pertinence des nœuds dans le cas de requêtes orientées contenu est calculé grâce à la propagation des poids des termes les plus spécifiques dans l'arbre du document. Les poids sont cependant diminués par multiplication par un facteur, nommé facteur "d'augmentation".

Par exemple, considérons la structure de document de la **figure 1.14**, contenant un certain nombre de termes pondérés (par leur probabilité d'apparition dans l'élément), et la requête "XML".

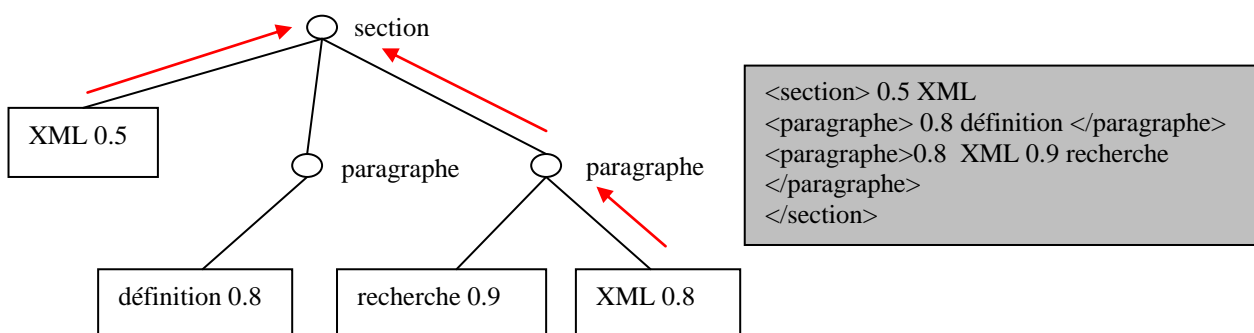


Figure 1. 14 : Modèle d'augmentation, [Fuhr *et al.*, 03]

Le poids de pertinence de l'élément section est calculé comme suit, en utilisant un facteur d'augmentation égal à 0.7 : $P([section, XML]) + P([paragraphe[2]], P([paragraphe[2], XML]) - P([section, XML]).P([paragraphe[2]]).P([paragraphe[2], XML) = 0.5 + 0.7*0.8 - 0.5*0.7*0.8=0.68$.

Le nœud paragraphe (ayant une pertinence de 0.8 à la requête) sera donc mieux classé que le nœud section.

Pour les requêtes orientées contenu et structure, des probabilités d'apparition de chaque terme de la condition de contenu dans les éléments répondant aux conditions de structure sont calculées, et des sommes pondérées de ces probabilités sont ensuite effectuées.

Nous trouverons une autre application du modèle probabiliste dans [Kompaoré, 08], avec une application de la théorie de Dempster-Shafer, [Shafer, 76]. La théorie de l'évidence de Dempster-Shafer est utilisée principalement parce qu'elle possède une règle de combinaison permettant d'effectuer une agrégation du score de pertinence des éléments en respectant la théorie de l'incertain.

II.1.6. Modèle inférentiel

Dans la recherche d'information dans des documents XML, les diagrammes d'inférence ont été adaptés pour exprimer les relations de causalité entre termes et structures.

Parmi les travaux de la littérature, citons celui Piworwarski *et al.* [Piworwarski *et al.*, 02]. Les auteurs ont proposé un modèle probabiliste basé sur les réseaux bayésiens où les dépendances de hiérarchisation sont exprimées par des probabilités conditionnelles. La probabilité de pertinence d'un élément e sachant son parent p pour une requête q est $P(e|p; q)$ est la suivante :

$$P(e = a|p = b, q) \cong \frac{1}{1 + e^{F_{e,a,b}(q)}} \quad (1.16)$$

où,

$F_{e,a,b}(q)$ est la pertinence de l'élément e selon le modèle Okapi.

Une requête q structurée est décomposée en un ensemble de n sous-requêtes élémentaires q_i . Chacune de ces sous-requêtes reflète une entité structurelle et un besoin d'information. Le score final est donné par la formule suivante :

$$RSV(e_i, q) = RSV_{q_1}(e_i, q) * \dots * RSV_{q_n}(e_i, q) \quad (1.17)$$

De Compos, Fernandez et Huete, [Fernandez-Luna *et al.*, 04], ont également proposé un modèle de recherche basé sur les réseaux bayésiens où le diagramme d'inférence est basé sur la probabilité conditionnelle. Deux types de diagrammes sont proposés : **S**imple **I**nference **D**iagram (SID) et **C**ontext based **I**nference **D**iagram (CID).

Un diagramme se compose de deux parties : une partie qualitative et une partie quantitative.

- Le composant qualitatif est la représentation des variables et des influences. Ceci est réalisé par les différents types de nœud : nœud de chance, nœud de décision et nœud d'utilité. Dans ce modèle les arcs pointent vers les nœuds de chance et d'utilité seulement.

Dans SID, les nœuds de chance et de décision sont liés aux nœuds de l'utilité qui seront additionnés à la fin. Dans CID, on rajoute par rapport à SID des arcs provenant des nœuds de chance vers les nœuds d'utilité qui leurs sont au dessus par rapport à l'arborescence.

- Le composant quantitatif est la probabilité des nœuds de chance et des nœuds d'utilité variant entre 0 et 1.

II.1.7. Le modèle différentiel bayésien

Les réseaux inférentiels bayésiens, [Turtle *et al.*, 90], considèrent le problème de la recherche d'information d'un point de vue épistémologique. Ils associent des variables aléatoires avec les termes de l'index, les documents et les requêtes de l'utilisateur. Les termes

de l'index et les documents sont représentés comme des noeuds. Une variable aléatoire associée avec un document d_j représente l'événement d'observer ce document. Les arcs sont dirigés du noeud document vers ses noeuds termes : ainsi, l'observation d'un document est la cause d'une augmentation de la valeur des variables associées avec ses termes d'index. La variable aléatoire associée à la requête de l'utilisateur modélise l'événement que la requête d'information spécifiée dans la requête a été vérifiée. La valeur de ce noeud requête est une fonction des valeurs des noeuds associés aux termes de la requête. Ainsi, les arcs sont orientés des noeuds des termes de l'index vers le noeud de la requête.

II.1.8. Les modèles de langage

Dans les modèles de recherche classique, on cherche à mesurer la similarité entre un document D_j et une requête Q ou à estimer la probabilité que le document réponde à la requête ($P(D_j=Q)$). L'hypothèse de base dans ces modèles est qu'un document n'est pertinent que s'il ressemble à la requête. Les modèles de langage sont basés sur une hypothèse différente : un utilisateur en interaction avec un système de recherche fournit une requête en pensant à un ou plusieurs documents qu'il souhaite retrouver. La requête est alors inférée par l'utilisateur à partir de ces documents. Un document n'est pertinent que si la requête utilisateur ressemble à celle inférée par le document. On cherche alors à estimer la probabilité que la requête soit inférée par le document $P(Q=D_j)$. En se basant sur ce principe d'indépendance des termes (l'apparition d'un terme n'influe pas la probabilité d'existence d'un autre terme dans le document ou dans la requête), $P(Q=D_j)$ peut être réécrite de manière simple en [Ponte *et al.*, 98] :

$$P(Q|D_j) = \prod_{i=1}^n P(T_i|D_j) \quad (1.18)$$

Où

n est le nombre de termes dans la requête et T_i est un terme de la requête, ($1 \leq i \leq n$).

Afin de pallier le problème des termes de la requête absents des documents, (ceci conduirait systématiquement à $P(Q=D_j)=0$), on combine deux modèles de langage : celui du document et celui de la collection.

Etant donné une requête composée des termes T_1, T_2, \dots, T_n , les documents sont ordonnés selon la mesure suivante, [Ponte *et al.*, 98] :

$$P\left(T_1, T_2, \dots, \frac{T_n}{D_j}\right) = \prod_{i=1}^n (1 - \lambda_i P(T_i) + \lambda_i P(T_i|D)) \quad (1.19)$$

Cette mesure est une combinaison linéaire du modèle de document et du modèle de contexte du document (la collection), où :

- $P(T_i/D)$ est la probabilité d'un terme important dans le modèle de document,
- $P(T_i)$ est la probabilité d'un terme dans le modèle de la collection

et

- λ_i est une constante.

II.2. Projection sémantique

Afin de prendre en compte la sémantique lors de l'évaluation, et donc lors de la génération d'un nouveau classement et sa comparaison avec les classements par défaut des SRI, Bouramoul associe à chaque terme de la requête l'ensemble des mots qui lui sont sémantiquement liés. L'idée est de projeter les termes de la requête sur les concepts de l'ontologie en utilisant les deux relations sémantiques : "synonymie" et "hyponymie" pour extraire les différents sens de la requête. Par la suite l'ensemble des concepts récupérés pour chaque terme sont utilisés en conjonction avec le terme lui-même lors de la pondération par le module de calcul. L'objectif est de favoriser un document qui contient des mots sémantiquement proches à ce que l'utilisateur cherche, même si ces mots n'existent pas comme termes dans la requête.

Bouramoul utilise, à cet effet, l'ontologie WordNet selon le principe suivant : au départ il accède à la partie de l'ontologie contenant les concepts et les relations sémantiques, ces derniers sont utilisés pour récupérer tous les synsets relatifs à chacun des termes de la requête. Ces synsets sont enfin utilisés pour la construction du vecteur sémantique qui contient pour chaque terme de la requête, les synonymes et les hypernoms appropriés. La **figure 1.15**, illustre le fonctionnement du module de projection sémantique, [Bouramoul, 11].

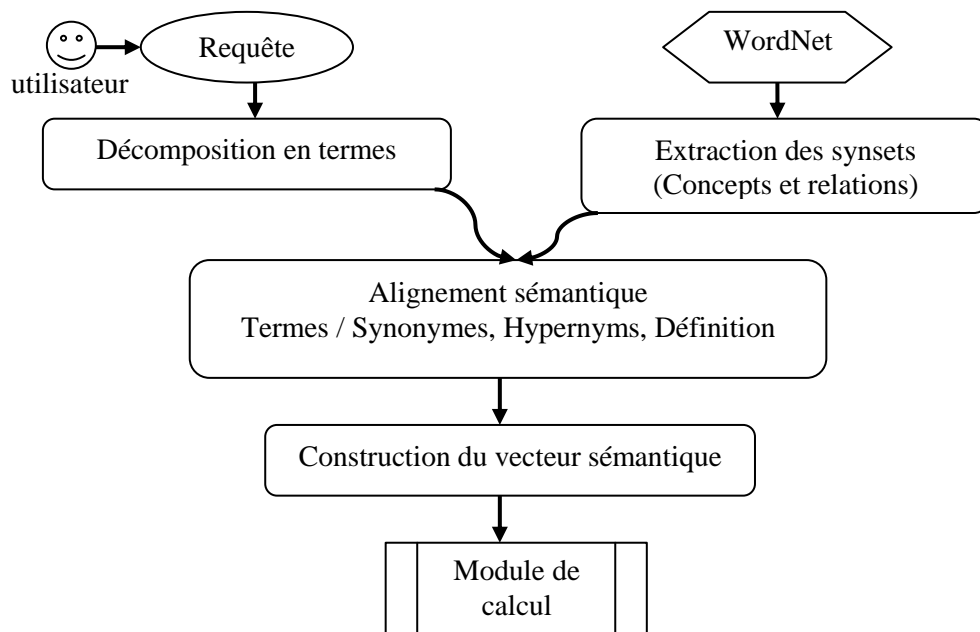


Figure 1.15 : Projection sémantique

II.3. Calcul de correspondance

Tout système de recherche d'information (SRI) s'appuie sur un modèle de recherche d'information. Ce modèle se base sur une fonction de correspondance qui met en relation les termes d'un document avec ceux d'une requête en établissant une relation d'égalité entre ces termes. Cette relation d'égalité représente la base de la fonction de correspondance et, par la même, du SRI, [Tambellini, 07] et [Audeh, 15].

Il existe un certain nombre de modèles théoriques dans la littérature les plus connus étant le « Modèle Booléen », le « Modèle Vectoriel » et le « Modèle Probabiliste ». Dans le modèle booléen, les requêtes sont représentés sous forme de termes reliés par des opérateurs booléens (ET, OU, NON, . . .). Le modèle vectoriel considère les documents et les requêtes comme des vecteurs pondérés, chaque élément du vecteur représentant le poids d'un terme dans la requête ou le document. Le modèle probabiliste tente d'estimer la probabilité qu'un document donné soit pertinent pour une requête donnée, [Kompaoré, 08].

II.4. Classement

Le rôle de la fonction de similarité est d'ordonner les documents avant de les renvoyer à l'utilisateur. En effet, l'utilisateur se contente généralement d'examiner les premiers documents retournés. Par conséquent, si les documents attendus ne sont pas présents dans cette tranche de résultats, l'utilisateur considérera le SRI par rapport à son besoin en information, et les résultats qu'il retourne seront donc considérés comme non pertinents. Dans ce contexte, le rôle de cette tâche est de finaliser le processus d'évaluation sémantique en rapprochant la pertinence système de la pertinence utilisateur.

À cette étape du processus d'évaluation, chaque document est décrit par deux valeurs de similarité générées par le module de calcul. En se basant sur la distance entre le vecteur document et celui de la requête, le module de classement procède à l'ordonnancement des résultats de façon à ce que le document ayant la plus faible valeur de distance, et donc la pertinence la plus élevée, sera classé premier jusqu'à ce que tous les résultats soient correctement arrangés.

Cette tâche prend également en charge la mesure de la pertinence de l'outil de recherche lui-même.

II.5. Présentation

Les résultats d'un SRI se présentent généralement sous la forme d'une liste de liens accompagnés d'un titre et d'un résumé décrivant le contenu de chaque page. Ces résultats, avant d'être présentés à l'utilisateur, doivent être ordonnés en fonction du score de pertinence attribué par les algorithmes propres à chaque SRI.

Dans l'approche proposée par Bouramoul, et afin de respecter ce principe généralement utilisé pour afficher les résultats d'une recherche, la tâche de présentation prend en charge la partie d'affichage une fois que les résultats sont traités. Plus précisément, cette tâche présente le compte rendu d'une session de recherche de la façon suivante :

- L'ensemble des résultats répondant à une requête, où chaque résultat est représenté par un triplet (titre, résumé, URL). Ces résultats sont classés sémantiquement selon le principe de l'approche proposée par Bouramoul.
- Le score de pertinence sémantique associé à chaque résultat.
- L'ensemble des concepts liés à chaque terme de la requête, ces concepts sont récupérés de l'ontologie WordNet et présentés sous forme d'une arborescence.

La **figure 1.16**, illustre la tâche de présentation, [Bouramoul, 11].

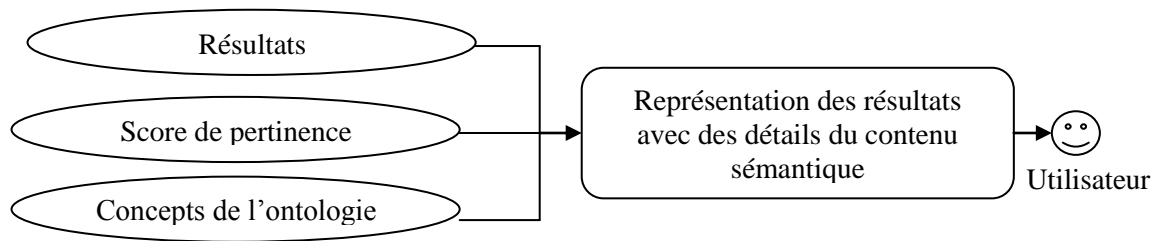


Figure 1. 16: Présentation

Conclusion

Les documents multimédia, en permettant le balisage des contenus des documents, réactualisent la problématique de recherche d'information classique, et permettent ainsi de traiter l'information avec une granularité plus fine. Le but des SRI traitant des documents multimédia est alors d'identifier des parties des documents les plus pertinentes à une requête donnée.

Nous avons ainsi présenté les principales approches d'appariement développées en SRI. Nous avons de plus indiqué que la reformulation de requêtes est une phase importante en RI classique permettant l'amélioration des performances des SRI.

Parmi les techniques les plus efficaces, nous citons la reformulation de requêtes par réinjection de pertinence. Avant de détailler notre contribution de trouver un SRI spécifique aux documents visioconférences médicales, nous allons présenter dans la partie suivante une brève discussion concernant l'état de l'art déjà présenté.

Discussion

De nos jours, nous assistons à un développement incessant des technologies de l'information. Ces nouvelles technologies ont permis l'évolution rapide de techniques matérielles de production et de gestion de l'information. Le progrès des outils de production d'information tels que les visioconférences a permis la production d'une énorme masse d'information. L'évolution des média électroniques a autorisé le stockage de cette vaste quantité d'information. Cette augmentation rapide du volume d'information a engendré le problème de « comment retrouver une information qui nous intéresse dans cette grande masse d'information ».

Afin de traiter ce problème, des SRI ont été développés pour sélectionner parmi un volume d'information, les informations pertinentes vis-à-vis d'un besoin en information. Ces SRI ont pour but de mettre en correspondance une représentation du besoin de l'utilisateur avec une représentation du contenu des documents au moyen d'une fonction de correspondance.

Généralement, un SRI se compose de trois modules : analyse, indexation et recherche. Les méthodes et outils présentés dans l'état de l'art montrent l'apport de l'utilisation d'une ontologie ou autres ressources utilisant un vocabulaire contrôlé, ainsi que les relations entre les termes afin d'indexer les documents. Le résultat de l'indexation n'est pas une liste de mots qui sont complètement indépendants mais qui sont reliés entre eux. L'apport de l'utilisation d'une ontologie est à deux niveaux :

- d'un côté, l'indexation des documents en utilisant un vocabulaire contrôlé et spécialisé à un domaine bien défini permet d'éviter d'avoir du bruit (trop de réponses dans le module de recherche) ;
- d'un autre côté, la recherche d'information en exploitant les relations entre les termes pour retrouver les documents les plus pertinents.

L'indexation sémantique à partir des ontologies est la méthode la plus pertinente et la plus prometteuse pour pallier aux problèmes de volatilité et d'hétérogénéité des documents, mais l'utilisation du contenu du document seulement peut générer un manque à cause des raisons suivantes :

- un auteur peut passer une idée dans un document sans utiliser les termes de l'ontologie ;
- et, un document peut contenir beaucoup d'idées sans avoir beaucoup de contenus. Par exemple un document décrivant les présentations d'une conférence va avoir peu de texte, mais de nombreuses références vers les présentations.

Pour conclure, nous avons présenté dans cette partie, le SRI avec ses trois modules. Le module d'indexation consiste à définir un ensemble de mots-clés pertinents décrivant le

document afin de tirer profit du contenu. Deux types d'indexations ont été retenus : l'indexation par le contenu et l'indexation par le contexte.

Le premier type consiste à indexer des documents à partir de leurs contenus, cette indexation peut se faire en utilisant ou non un vocabulaire contrôlé.

Nous avons vu dans cette partie l'utilité de suivre une ontologie ou un langage commun afin d'indexer les documents. Les ontologies servent dans le module d'indexation à plusieurs niveaux :

- l'utilisation d'un vocabulaire contrôlé facilite le partage des informations ;
- la désambiguïsation des termes lors de l'indexation, en se basant sur les cooccurrences des termes ;
- l'indexation des documents indépendamment de la langue, une ontologie ayant un ensemble de concepts dénotés par des termes dans plusieurs langues.

Le deuxième type d'indexation est par le contexte. Pour cela nous avons introduit quelques travaux correspondants à ce type d'indexation.

Étant dans un contexte technique, spécialisé et multimodales, nous utilisons des ontologies afin d'effectuer l'indexation des documents.

Le système que nous allons construire se compose de trois modules, des ontologies, d'une base de documents ainsi qu'une base de données index. Le module d'analyse (nommé aussi module de modélisation) se charge de parcourir et d'analyser les documents qui se trouvent dans la base de documents à l'aide du formalisme des graphes conceptuels. Il s'agit de définir un modèle pour la description du contenu des média qui sont appropriés à la représentation du contenu des visioconférences. Ce modèle doit prendre en compte les éléments d'information à différents niveaux de descriptions et les différents média présents et qui sont adaptés pour être intégrés dans un système d'indexation et de recherche. Le résultat de l'analyse est un fichier XML.

Le module d'indexation se charge de parcourir et d'indexer les documents résultats du module d'analyse à l'aide des concepts de l'ontologie. Le résultat de l'indexation est sauvegardé dans une base de données index. L'indexation d'un document se fait par l'administrateur du système qui soumet l'adresse du document à indexer au module d'indexation.

Dans le cas où un utilisateur veut chercher, une visioconférence donnée, sa requête sera envoyée au module de recherche. Ce dernier se compose essentiellement de deux étapes : la première est la reformulation de la requête initiale de l'utilisateur afin de produire une nouvelle requête qui reflète mieux son besoin. Dans cette étape, nous exploitons les liens sémantiques entre les concepts en utilisant une ontologie de domaine et le profil utilisateur pour élargir la requête. La deuxième étape est l'appariement requête-document. Elle se base sur une fonction de correspondance ou de ranking et permet de comparer la représentation de la requête à celle de chaque document du corpus. La fonction de correspondance consiste à calculer la similarité entre la représentation de la requête et de chaque document.

L'appariement requête-document est intimement lié au module d'indexation et de pondération des termes des requêtes et des documents. En effet, leurs représentations ainsi que leurs appariements permettant de définir un modèle de recherche d'information.

Les ontologies ont amélioré considérablement la pertinence des résultats dans la recherche dans des bases de documents audiovisuels comme déjà détaillé dans le deuxième chapitre. Elles diminuent le silence et le bruit. C'est la raison pour laquelle nous opterons pour la méthode de modélisation à l'aide de l'ontologie. Cette amélioration est due au fait que le module de modélisation prend en compte les concepts ainsi que les différentes relations entre eux (données par l'ontologie). Par conséquent, et contrairement aux méthodes par mots clés simples et statiques qui s'intéressent à l'existence ou non d'un mot dans un document, cette méthode prend en compte la sémantique des termes recherchés.

La partie suivante de notre thèse (notre contribution) est consacrée aux techniques et travaux sur la modélisation multi-facette et multimodale des documents multimédia et plus précisément un document visioconférence. Chaque facette correspond à un aspect de la représentation des documents pour en permettre la recherche efficace et pertinente. Ainsi, elle décrit le module d'indexation des documents visioconférences complété par l'utilisation des ontologies de domaine. Finalement, elle détermine l'appariement entre les requêtes utilisateur et l'index obtenu lors du module d'indexation, pour satisfaire les besoins de l'utilisateur après avoir reformulé la requête initiale de l'utilisateur.

Partie 2

*Contribution: OSSVIRI -
Ontology-based System for
Semantic Visio-conference
Information Retrieval and
Indexation*

Chapitre 1 : Modélisation structurelle et sémantique

La modélisation des connaissances est un sujet majeur pour différentes problématiques. Différentes méthodes, formalismes et outils ont été proposés et mis en œuvre pour modéliser, stocker et utiliser les informations au sein des bases de données ou des bases de connaissances pour répondre aux besoins d'un utilisateur jouant par la suite un rôle central dans l'interaction et le dialogue avec l'utilisateur tels que les graphes conceptuels et les ontologies.

Pour fournir plusieurs niveaux d'abstraction en exploitation du contenu visioconférence, nous proposons un schéma de modélisation à deux niveaux : niveau structurel et niveau sémantique. Les descriptions du niveau structurel sont calculées automatiquement en utilisant des descripteurs visuels. La modélisation sémantique est une abstraction qui permet de lier des descriptions de bas niveau avec le monde réel.

Introduction	74
I.Schéma de modélisation du contenu visioconférence	74
II.Modélisation structurelle : multi-facette	78
II.1. La notion de facette (ou vue)	78
II.2. Modélisation du contenu visuel	78
II.2.1. La facette perceptive	78
II.2.2. La facette structurelle	79
II.2.3. La facette spatiale.....	79
II.2.4. La facette symbolique	80
II.2.5. Facette événementielle	82
II.2.6. Facette temporelle	83
II.2.7. Facette signal	85
II.3. Modélisation du contenu audio	94
II.3.1. Spécification conceptuelle et description du contenu audio	94
II.3.2. Segmentation audio et structure du document	95
II.3.3 Modélisation de la sous-facette audio	95
III. Modélisation sémantique	97
III.1. Modélisation du contenu visuel : Facette sémantique	97
III.2. Modélisation du contenu Audio : extraction des termes	97
III.2.1. Transcription de la parole	98
III.2.2. Extraction des termes simples.....	102
III.2.3. Extraction des Termes composes.....	106
Conclusion.....	110

Introduction

Notre sujet de thèse, comme c'est déjà cité, vise à rechercher des documents visioconférences dans des bases de données de grande taille tout en permettant l'accès de l'utilisateur à l'information qu'il cherche en facilitant l'exploration de ces documents à partir des requêtes relatives à leurs contenus. Et pour que ces visioconférences soient exploitables, il faut qu'elles puissent être consultées efficacement. En effet, une indexation, se propose d'attacher à ces documents audiovisuels un ensemble de descripteurs de leur contenu, dans le but de mesurer la ressemblance avec les descripteurs correspondants à la requête. Cette requête constitue l'expression du besoin en information de l'utilisateur.

Donc, l'indexation a pour objectif non pas d'identifier les informations contenues dans les documents mais plutôt de caractériser les connaissances associées à ces documents. Elle consiste par conséquent à extraire des informations synthétisant leurs contenus et à récupérer tout ou partie d'un document original. Elle doit être exploitable c'est-à-dire le document indexé doit être mis en valeur pour une réutilisation maximale et opportune à moindre effort d'indexation. Aussi, elle doit être concise et automatique. Pour achever ces buts, le langage d'indexation doit être le plus standard possible, et l'expression sémantique doit être la plus partagée possible. Toutefois, en l'absence d'outil entièrement automatique, l'indexation est lourde en mobilisation de ressources humaines. En outre, elle doit être lisible. Cet aspect (la lisibilité) est peut être amélioré à l'aide d'outils de représentation textuelle ou graphique.

Les visioconférences ont généralement un but pédagogique. Et vu qu'un document pédagogique puisse être réutilisé dans différents contextes (ce qui est le but recherché), l'indexation doit en favoriser une utilisation optimale en fonction d'un besoin utilisateur précis ou d'une application précise.

Avant la phase d'indexation, il faut passer par une phase de modélisation. En fait, l'enjeu de la représentation des connaissances est de permettre d'explicitier des connaissances humaines de toutes sortes dans un formalisme interprétable par une machine. Un formalisme de représentation doit être suffisamment expressif pour que l'homme puisse aisément modéliser et modifier toutes les connaissances utiles dans un domaine d'application donné.

Nous détaillerons, dans ce chapitre, étant donné que l'objectif primordial consiste en la conception d'un SRI pour les visioconférences basé sur les graphes conceptuels, les différents points reliés au modèle que nous avons proposé. Notre travail s'inscrit dans un contexte d'indexation et de recherche des visioconférences par le contenu sémantique. Nous commençons par définir le formalisme de modélisation. Nous décrivons dans un premier lieu notre modélisation générale du contenu d'une visioconférence et dans un deuxième lieu la modélisation de chaque type de média. Enfin, nous décrivons notre modèle via un exemple.

I. Structure générale de notre SRI

Les trois modules de notre système proposé sont illustrés dans la **figure 2.1**.

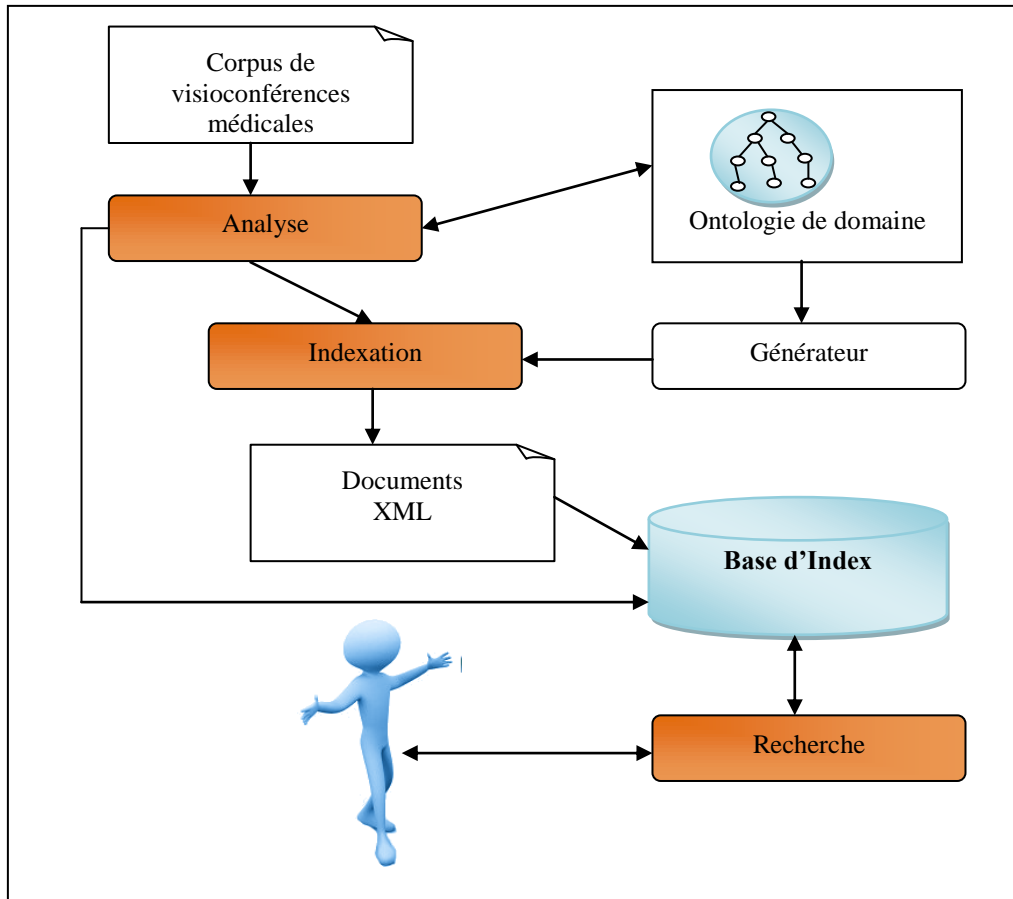


Figure 2. 1: Architecture générale de notre système

II. Schéma de modélisation du contenu visioconférence

L'information du contenu vidéo peut être représentée sous plusieurs niveaux : les informations physiques constituées des données binaires du contenu qui n'est pas utilisable que par l'ordinateur, et les informations de description qui permettent de transformer les informations physiques en connaissances exploitables par l'utilisateur, ce qui permet de renforcer l'interface entre l'homme et la machine et d'exploiter donc plus facilement le contenu vidéo. Pour fournir plusieurs niveaux d'abstraction en exploitation du contenu vidéo, nous proposons un schéma de modélisation à deux niveaux : structure et sémantique.

La modélisation de la structure du contenu permet de décrire l'organisation qui peut représenter l'information du contenu. Cette modélisation est souvent basée sur la structure classique d'un document vidéo (la séquence, la scène, le plan). Les descriptions de ce niveau sont calculées automatiquement en utilisant des descripteurs visuels. Ce niveau de modélisation est dépourvu de toute description sémantique, [Tapu, 12].

La modélisation sémantique est une abstraction qui permet de lier des descriptions de bas niveau avec le monde réel. Un schéma de modélisation qui permet de décrire la signification des descriptions situées au niveau de la structure. Nous exploitons la notion des concepts et

des relations conceptuelles pour présenter les occurrences (éléments d'information) décrites dans la partie structure.

Notre approche de modélisation se situe au niveau de la description conceptuelle. Il s'agit de concevoir un schéma de représentation générique indépendant du contenu vidéo mais qui permet de renseigner les différents éléments d'informations se situant au niveau sémantique (cf. **Figure 2.2**).

Si nous voulons modéliser le contenu d'une visioconférence, nous pensons à deux types de modélisation qui sont modélisation de sa structure et modélisation sémantique de son contenu.

Pour la modélisation de la structure, les visioconférences sont hiérarchiquement structurées en plans et images clefs. Une telle structure reflète le processus de création de la visioconférence.

Dans une visioconférence, nous pouvons avoir plusieurs emplacements pour le déroulement des différentes actions. Par exemple dans une visioconférence médicale, les événements peuvent se dérouler dans la salle d'opération, dans la salle des réunions, en dehors de ces dernières salles... etc. Nous pouvons à chaque moment avoir une intervention d'un nouveau participant si nous avons des éventuels points de vue de différents experts ou participants. Ceci montre qu'elle est formée des plans.

La différence entre la structure d'une vidéo et celle d'une visioconférence est qu'une vidéo peut être constituée par plusieurs scènes. Une scène est définie comme une suite de plans contigus pouvant être sémantiquement reliés c'est à dire ayant une cohérence sémantique, alors que la totalité de la visioconférence constitue une seule scène puisque depuis son début jusqu'à sa fin elle évoque le même thème.

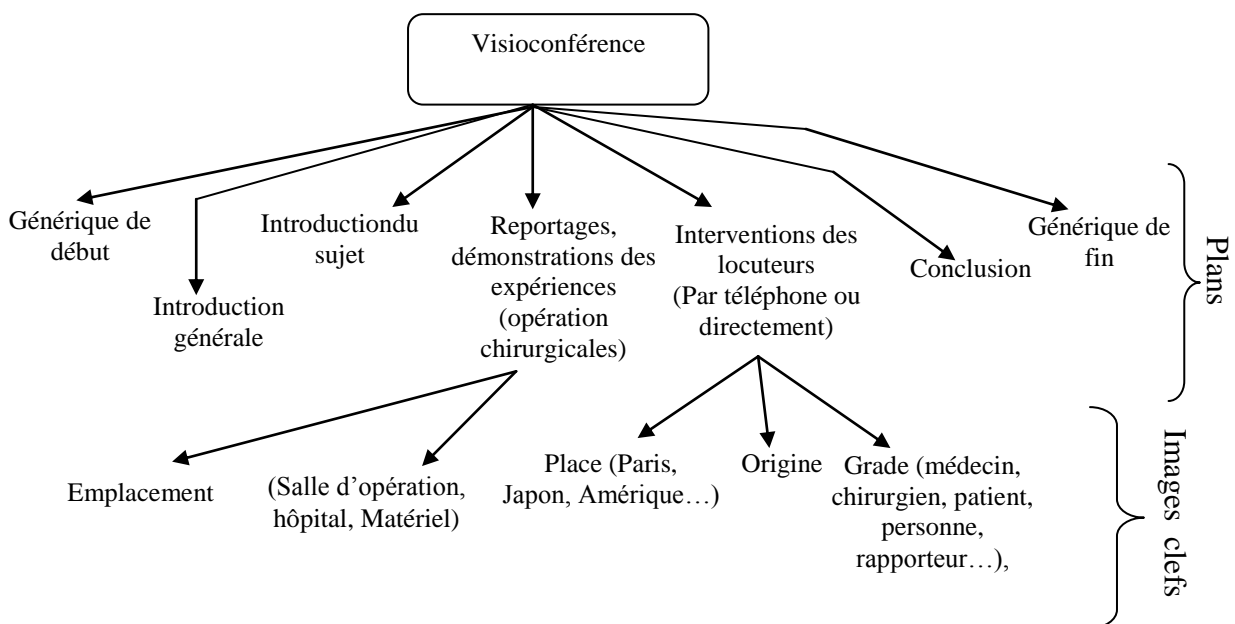


Figure 2. 2: Structure d'une visioconférence

Les plans sont définis comme des ensembles d'images tournées sans arrêt de la caméra. Un plan est considéré comme l'unité de montage vidéo (notion d'image(s) clé (keyframe)). Les scènes sont définies comme des suites de plans contigus pouvant être sémantiquement reliés c'est à dire ayant une cohérence sémantique. Pour dégager cette organisation hiérarchique, nous s'intéressons aux descripteurs visuels seulement. Elle se définit selon une approche descendante par raffinements successifs. Pour un utilisateur, le fait de parcourir une structure hiérarchique, est certainement plus aisé que de naviguer sur une structure plate.

L'intérêt principal de cette organisation est qu'elle peut être extraite automatiquement en s'affranchissant de la sémantique du contenu. Quant à la modélisation sémantique, nous allons utiliser le formalisme des graphes conceptuels pour présenter les différentes idées figurant dans une visioconférence en utilisant des concepts et des relations. Il est recommandé de décrire le contenu de chacun des éléments structurels afin d'extraire des extraits de la visioconférence répondant à des requêtes spécifiques en navigant dans les graphes conceptuels.

Le modèle des graphes conceptuels est une approche de modélisation ayant les spécificités d'être formelle, de représenter des connaissances, et d'être concrète dans le sens où il existe des outils efficaces pour manipuler les connaissances modélisées.

Il permet de modéliser les connaissances d'un domaine en termes de graphes, basés sur un support. Cette approche de modélisation est intensionnelle, est munie d'une sémantique en logique du premier ordre, et fait l'hypothèse d'un monde fermé pour ses raisonnements, **[Djamel et al., 08]**.

Dans notre approche de modélisation, nous allons concevoir un schéma de modélisation des connaissances sémantiques combinant les deux types de modélisation (hiérarchique et sémantique) et indépendamment du contenu de la visioconférence c'est-à-dire que nous pouvons l'appliquer sur toutes les visioconférences.

Les descripteurs ou caractéristiques multimédia sont des données qui décrivent elles-mêmes le contenu des données mais de façon beaucoup plus concise. Ces descripteurs consistent à représenter sous une forme réduite le contenu des documents et suscitent un besoin de normalisation du vocabulaire.

D'une façon générale, la plupart des algorithmes traitent un seul média, dans l'analyse de contenus multimédia, puisque les logiciels de recherche sur le Web n'utilisent que des descriptions textuelles et n'utilisent ni l'image ni l'audio. Faut-il signaler que les systèmes de recherche d'images mélangent les descriptions visuelles et textuelles.

Ceci est encore le cas pour l'analyse des vidéos qui sont habituellement faites en séparant les séquences sonores et l'image. Les algorithmes qui sont spécialisés sur le traitement d'un médium peuvent être plus rapides qu'un autre qui indexe tous types de contenu.

Nous cherchons dans notre approche à rendre l'indexation des documents visioconférences plus efficace tout en prenant en compte le fait qu'une visioconférence intègre de sources hétérogènes (texte, image, audio). Ces sources coopèrent pour répondre à une requête de

l'utilisateur. Chaque source étant en mesure de fournir une partie des réponses ou encore des réponses partielles.

III. Modélisation structurelle : multi-facette

III.1. La notion de facette (ou vue)

À chaque facette, définissant la correspondance avec un type du contenu informatif sur l'ensemble des objets images, est associé un modèle décrivant les objets images, les relations qui les lient et les opérations définies sur ces descriptions. Les facettes spécifiées sont complémentaires dans un souci de correspondance avec les interprétations qu'elles modélisent et chaque instance du modèle général est une combinaison de facettes pour traduire la richesse des caractéristiques pertinentes des images, [Charhad, 05].

Deux catégories de facettes principales se distinguent : la facette physique qui représente l'entité perçue par l'œil humain dans sa représentation plane et bidimensionnelle, et la facette logique rassemblant les interprétations de l'image et l'ensemble de ses descripteurs plus sémantiques. Cette dernière se subdivise elle-même en quatre facettes dont la combinaison fournit la caractérisation symbolique de l'image : les facettes structurelles, spatiales, symboliques et perceptives, [Soulier *et al.*, 13] et [Zwidi *et al.*, 13].

III.2. Modélisation du contenu visuel

Une image physique est subdivisée en un ensemble de régions représentant son contenu. Une région est par conséquent elle-même un objet physique véhiculant une information indispensable pour la modélisation du contenu global de l'image.

Chaque région est représentée par une abstraction dite « Objet image » (noté I_o) dont le référent physique est la région elle-même, [Charhad, 05].

Pour modéliser le contenu visuel d'une visioconférence, nous travaillons d'une manière analogue celui du modèle EMIR² combinant différentes interprétations (facettes) de l'image pour construire une description complète de son contenu. Ce modèle permet une représentation symbolique et multi-facettes du contenu sémantique des images.

III.2.1. La facette perceptive

La facette perceptive regroupe l'ensemble des attributs visuels associés à l'image entière mais aussi aux différents objets image tels que la couleur, la texture, la forme etc. Un des principaux avantages dans l'utilisation de ces attributs réside dans leur caractère objectif est dans la possibilité de les calculer de façon automatique.

Dans EMIR², trois attributs sont pris en compte : la couleur, la texture et la brillance.

- la couleur : peut être défini par rapport à un espace de couleur particulier. Il existe plusieurs espaces de couleur (RVB, CMYk, ...)
- la texture : elle est définie comme le graphique qui couvre la surface d'une image ou d'un objet ;
- la brillance : elle mesure la luminosité moyenne de l'image ou d'un objet image.

III.2.2. La facette structurelle

La facette structurelle représente la décomposition d'une image en objets images. Chaque objet image peut également être décomposé en sous-objets images. La relation de composition associée à cette facette est la relation « contient » qui implique l'inclusion spatiale, les régions correspondant aux objets composants sont incluses dans les limites géométriques de la région décrite par l'objet décomposé.

La facette structurelle est représentée par un graphe conceptuel dont les nœuds sont les objets images et les arcs les instances de la relation de composition.

Le modèle de la facette structurelle est défini comme le couple (Ioi, CONP) avec :

- Ioi : ensemble des objets images de la facette structurelle.
- $CONP \subseteq Ioi \times Ioi$: relation de composition entre les objets images.

La facette structurelle introduit un seul type d'objet (ObjetImage) et un seul type de relation (COMP) formant ces treillis de types. Le graphe d'une vue structurelle est construit en liant des objets de type ObjetImage et dont les référents sont les identifiants d'objets images par la relation conceptuelle de type COMP. Le graphe de base générant par jointure les graphes des facettes structurelles est le suivant :

$$[ObjetImage] \rightarrow (COMP) \rightarrow [ObjetImage]$$

III.2.3. La facette spatiale

La facette spatiale décrit les informations géométriques relatives aux objets spatiaux associés aux objets images ainsi que les relations spatiales entre eux. Cette facette permet de caractériser les objets d'une image par leurs formes et leurs positions relatives. Un objet spatial est défini par la donnée d'une forme géométrique (point, segment, polygone) correspondant à son contour. La facette spatiale est représentée dans les espaces de classiques afin de conférer au modèle une plus grande généralité. Ces espaces sont tout d'abord l'espace euclidien regroupant les notions de produit scalaire, orthogonalité, angles et normes. Cet espace permet les opérations telles que le calcul du barycentre, de la surface, de la longueur, de la largeur, de la hauteur et du polygone englobant, [Zwidi *et al.*, 13].

L'espace topologique exprime les notions de continuité et de connexion illustrés par les relations Couvre, Touche, Coupe, Dans, Disjoint, Intersection. Cet ensemble de six relations regroupe tous les cas de relation spatiale entre deux objets est invariant par rapport aux transformations géométriques de base.

L'espace vectoriel regroupe les relations de direction fondées sur des notions géographiques (« Nord », « Sud », « Ouest », « Est »). Enfin l'espace métrique définit les opérations permettant le calcul des caractéristiques géométriques des objets telles que la distance, la distance minimale, la distance normalisée et la distance maximale ainsi que les relations les situant les uns par rapport aux autres. Celles-ci sont au nombre de deux : « Proche » et « Loin » et sont pondérées par une valeur de l'intervalle [0,1] qui mesure le degré d'éloignement des objets spatiaux considérés, [Yengui *et al.*, 09].

Le modèle de la facette spatiale est défini par le 6-uplet (Isp, POINT, OS, RSPA, forme, Rsp) avec :

- Isp : ensemble des identificateurs des objets images de la facette spatiale ;
- POINT = $N \times N$: ensemble des coordonnées de points ;
- OS : ensemble des objets géométriques de base pour représenter la forme des objets ;
- RSPA = {loin, près, est, ouest, nord, sud, dans, disjoint, touche, couvre, coupe, ...} : ensemble des relations spatiales ;
- forme : $Isp \rightarrow P(OS)$ associe à chaque identifiant des objets spatiaux aux objets géométriques de base ;
- Rsp $\subseteq RSPA \times Isp \times Isp$: ensemble des relations spatiales qui lient les objets spatiaux dans l'image.

Les types d'objets géométriques représentant les formes possibles pour les objets spatiaux sont représentées par des types d'objets et organisés dans un treillis. Les relations spatiales sont organisées en treillis. Les objets spatiaux et leur forme sont représentés par des objets sous-types d'Objet-Spatial dont les référents sont leurs identifiants.

L'instance de la facette spatiale d'une image est alors représentée par un graphe conceptuel composé des concepts correspondants aux objets spatiaux et qui sont liés par les relations spatiales.

III.2.4. La facette symbolique

La facette symbolique correspond à la représentation du contenu sémantique d'une image et se définit comme la donnée d'objets symboliques associés aux objets images ainsi que par des relations correspondant à la description de scènes ou d'actions faisant intervenir ces objets. La facette symbolique s'efforce de prendre en compte les interprétations multiples quant à la sémantique véhiculée par l'image. Elle est fortement contrainte par l'application dans la mesure où l'expression du « sens » est liée à la connaissance du domaine de l'application ainsi que d'un langage d'indexation choisi pour exprimer les relations entre les éléments de connaissance mis en lumière.

Dans *EMIR*², un objet symbolique est décrit par un attribut via un treillis de classes muni d'une relation d'ordre partiel noté \leq , [Charhad, 05]. Il est aussi exprimé par un ensemble de couples « propriété, valeur », le champ propriété correspondant à des propriétés telles que le nom d'un patient et son numéro de dossier dans le contexte des applications médicales.

Les relations symboliques sont des relations binaires associées à des couples de classes de descripteurs $R(c_1, c_2)$. R décrit donc dans une image une relation entre les objets symboliques o_1 et o_2 appartenant respectivement aux classes cl_1 et cl_2 si et seulement si $cl_1 \leq c_1$ et $cl_2 \leq c_2$.

La facette symbolique est définie par rapport au modèle sémantique d'une application. Celui-ci regroupe les classes d'objets et leur treillis, la relation de composition entre ces classes, les relations symboliques et les propriétés, [Zwidi *et al.*, 13].

Le modèle sémantique d'une application M est défini par le 7-uplet (IDcl, IDpr, IDrs, VAL_PROP, PROP, RSYMB, domaine) avec :

- IDcl : ensemble des identificateurs des classes d'objets (relation d'ordre partielle, élément minimal m et élément maximal M). Chaque classe de IDcl est représentée par un type de concept. La relation d'ordre partiel entre les classes d'objets images est représentée dans le treillis des types de concepts.
- IDpr : ensemble des identificateurs des propriétés représentés par des types de relations conceptuelles dans le treillis des types de relations.
- IDrs : ensemble des identificateurs des relations symboliques représentés par des types de relations conceptuelles dans le treillis des types de relations.
- VAL_PROP = Réel U Entier U Chaîne U Booléen : ensemble des définitions de propriétés.
- PROP \subseteq IDpr \times IDcl \times P(VAL_PROP) : ensemble des définitions de propriétés représentées par le graphe canonique $[c] \rightarrow [p] \rightarrow [val]$ où p est la relation conceptuelle correspondant à la propriété p de IDpr.
- RSYMB \subseteq IDrs \times IDcl \times IDcl : ensemble des définitions des relations symboliques représentées par le graphe canonique $[c1] \rightarrow [rs] \rightarrow [c2]$ où rs est la relation conceptuelle correspondant à la relation rs de IDrs.
- Domaine : IDpr \rightarrow P(VAL_PROP) définit pour chaque type de valeurs de propriétés l'ensemble des valeurs possibles.

La facette symbolique associée à l'ensemble des objets symboliques leur sémantique définie dans le modèle d'application.

Le modèle de facette symbolique est défini comme le 5-uplet (M, Isy, cl, RI, PI) avec :

- M : application pour laquelle le contexte symbolique est défini.
- Isy : ensemble des identificateurs d'objets symboliques.
- cl : Isy \rightarrow IDcl associe à chaque objet symbolique sa classe.
- RI \subseteq IDrs \times Isy \times Isy : lie les objets symboliques entre eux par les relations de IDrs.
- PI \subseteq IDpr \times Isy \times VAL_PROP : lie les objets images à leurs propriétés.

Dans la représentation de la facette symbolique dans le formalisme des graphes conceptuels, le modèle de la facette symbolique est confondu avec celui du modèle de l'application. Il comprend la définition du type de concept représentant les objets symboliques, l'ensemble des classes et leur relation d'ordre, la définition des propriétés, leurs valeurs et la relation de composition entre les classes d'objets.

Une instance de facette symbolique est représentée par un graphe conceptuel canonique produit par la jointure sur le concept objet symbolique entre les graphes correspondant aux propriétés et ceux correspondant aux objets symboliques liés par des relations symboliques.

III.2.5. Facette événementielle

Par définition, un événement est un fait qui survient à un moment donné. Au sens général, il signifie tout ce qui arrive et possède un caractère spécifique. Dans le cas de la vidéo, un événement est quelque chose qui arrive dans un document ou dans un segment de document vidéo, [Habibian *et al.*, 14].

Dans notre proposition, nous considérons la modélisation du contenu visioconférence basée sur la description des événements comme étant une modélisation générique (indépendant d'un médium particulier). En effet, généralement même si un événement peut être décrit d'une manière spécifique au type de média (on entend le bruit d'une explosion par exemple), au niveau interprétation la description d'un événement qui est associée au segment vidéo et non au type de média. Un événement peut correspondre à une interaction entre les entités audiovisuelles (personne, objets, etc.).

III.2.5.1. Modélisation de la facette événementielle

Considérons $Ev = \{Ve1, Ve2, Ve3, Ve4, \dots, Ven\}$ l'ensemble des événements. Un événement peut contenir des sous-événements (qui sont aussi des événements).

Un événement peut s'étendre sur un document complet (événement global). Ce type d'événement est généralement décomposable en sous-événements (*cf.* **Figure 2.3**). Il en apparaît souvent dans le cas des documents spécifiques tels que les opérations chirurgicales. En effet, une opération chirurgicale peut constituer un événement et une anesthésie dans cette opération est alors considérée comme un sous-événement dans ce même document, [Mazloom *et al.*, 13].

Afin de modéliser l'ensemble des événements et des sous-événements dans un document (ou un segment de document) visioconférence, nous utilisons une relation de composition « est sous-événement de » qui permet de décomposer un événement en des sous-événements qui peuvent être eux-mêmes décomposables ou élémentaires. Un événement est élémentaire lorsqu'il constitue une description événementielle indécomposable.

Dans la **figure 2.3.**, l'événement $Evt3$ correspond à un sous-événement de l'événement $Evt1$ dans le contexte du document visioconférence. Il peut être une simple action dans le document. Une action peut être « filmique » telle qu'un mouvement de la caméra pour un spécialiste de montage visioconférence ou bien elle peut être « réelle » dans la visioconférence comme le mouvement des acteurs dans un film par exemple.

Comme le décrit le treillis représenté dans la **figure 2.3**, dans notre proposition pour la modélisation des « événements » dans la visioconférence, nous nous limitons à une description des événements représentant les actions réelles dans le document telles que par exemple action d'une personne (exemple : un discours) ou bien action d'un groupe de personne (exemple : une manifestation). Notons que d'une manière générale, la construction de ce type de treillis ainsi que la classification des événements dépend de l'application. D'autres événements un peu plus génériques sont classés selon un sujet particulier dans le document comme par exemple des événements chirurgicaux.

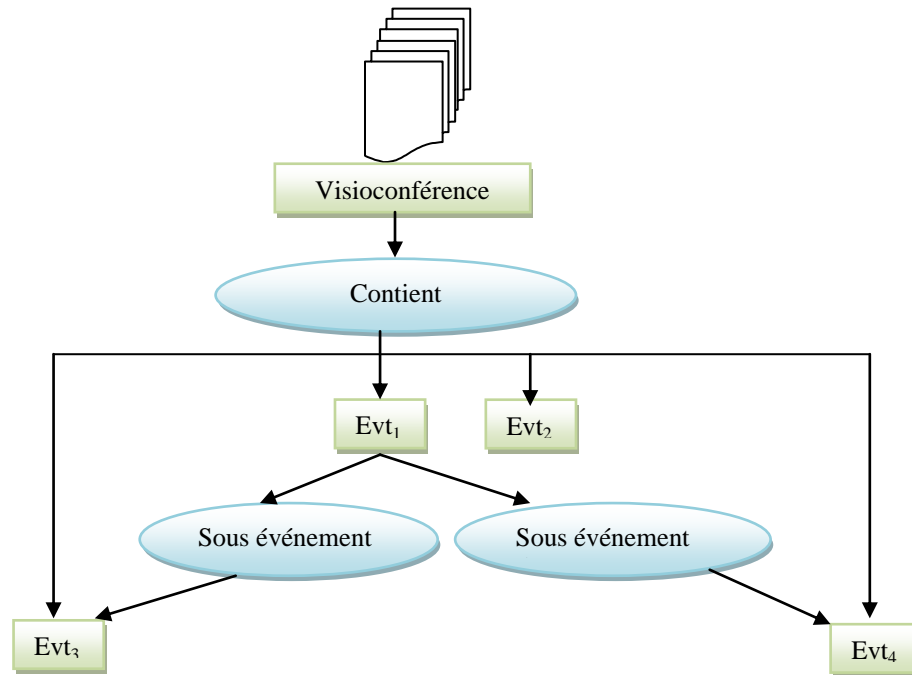


Figure 2. 3: Description événementielle

III.2.5.2. Spécification conceptuelle

La spécification conceptuelle est la description des différentes actions que nous pouvons l'extraire d'un document visioconférence. Chaque description est définie par l'ensemble des faits se produisant dans le passage. Les descriptions événementielles sont représentées par les objets « événement » **Ve**. Chaque représentation est formulée par un ensemble de graphes conceptuels comme pour le cas de la modélisation du contenu visuel.

La génération basique de graphe de représentation est la suivante :

$$[Ve1] \rightarrow (\text{sous-événement de}) \rightarrow [Ve2]$$

La description des événements dans un document vidéo s'attache aussi à la nature temporelle du document. En effet, les événements ont une extension temporelle qui correspond à un intervalle, ce qui permet de définir des relations temporelles entre eux. Les relations temporelles permettent aussi de décrire les objets (les apparitions / disparitions de ces objets), (*cf.* **Figure 2.3**).

III.2.6. Facette temporelle

L'information temporelle est une caractéristique spécifique du document visioconférence qui permet de synchroniser et d'ordonner l'ensemble des descriptions utilisées pour la modélisation du contenu visioconférence. En effet, c'est elle, en donnant au document son caractère dynamique, qui le différencie d'un document classique.

Des nombreux de travaux de recherche ont centré leurs études sur l'exploitation de l'information temporelle dans le document visioconférence. Dans cette section, nous n'avons pas l'objectif de ré-décrire les approches proposées dans ces travaux. Nous allons plutôt

détailler comment nous avons exploité et intégré cette caractéristique temporelle dans notre modèle.

Pour intégrer l'information temporelle dans le modèle, nous proposons de représenter le contenu d'une séquence par deux descriptions conjointes. La première modélise le contenu sémantique pour chaque **Unité Audio Visuelle** (UAV définit une entité abstraite représentant un segment quelconque de document vidéo) présente dans la séquence. La seconde modélise l'information temporelle entre les différentes UAV décrites dans la séquence.

Les relations d'Allen, [Hjelsvold *et al.*, 94], constituent un bon exemple pour décrire le contenu de la vidéo en se basant sur l'information temporelle. Ces relations sont explicitées dans la **figure 2.4**. Sur cette figure, nous illustrons les relations temporelles qui peuvent exister entre deux UAV. Il existe au total 13 relations d'Allen dont douze relations sont asymétriques. Excepté la relation « égale », les autres peuvent être regroupées deux par deux.

En effet, si nous avons UAV₁ chevauche UAV₂ alors nous pouvons inférer le que UAV₂ est chevauchée par UAV₁. C'est pour cette raison dans notre modèle nous ne gardons que sept relations.

L'information temporelle est une caractéristique spécifique à la visioconférence. Elle peut être associée dans le flux visuel à l'apparition ou à la disparition des objets visuels par exemple ou bien dans le contenu audio à un changement de thème dans le discours ou à un changement de locuteur.

La segmentation du contenu audio ne correspond pas forcément à un segment issu du contenu visuel d'un point de vue de la durée et des bornes des intervalles. Notons que la segmentation du contenu audio ne correspond pas forcément à celle du contenu visuel.

III.2.6.1. Modélisation de la facette temporelle

La facette temporelle contient l'ensemble des relations qui relient les entités d'une visioconférence dans un ordre temporel bien déterminé.

Afin de modéliser des descriptions temporelles, nous considérons d'abord un ensemble des relations temporelles ; sept relations qui sont appropriées pour la description temporelle sont choisies. Ces relations sont décrites entre UAV (*cf. Figure 2.4*).

III.2.6.2. Spécification conceptuelle

Les événements sont reliés deux à deux par des **relations temporelles** (TeR). Par exemple, Ve₁ et Ve₂ sont reliés par TeR, est traduit comme UAV₁ se déroule avant UAV₂ (UAV₂ après UAV₁).

La génération de base pour décrire l'ensemble des graphes attachés à cette facette entre deux unités audiovisuelles est illustré dans la représentation linéaire par :

$$[UAV1] \rightarrow (TeR) \rightarrow [UAV2].$$

Notons que pour cette facette temporelle, notre modélisation est restreinte uniquement à une description par des relations symboliques. Elle ne permet pas de prendre en compte des informations de type « timecode » et durée. Ceci est du principalement à l'utilisation des

formalismes des graphes conceptuels qui ne permet pas de représenter des valeurs numériques.

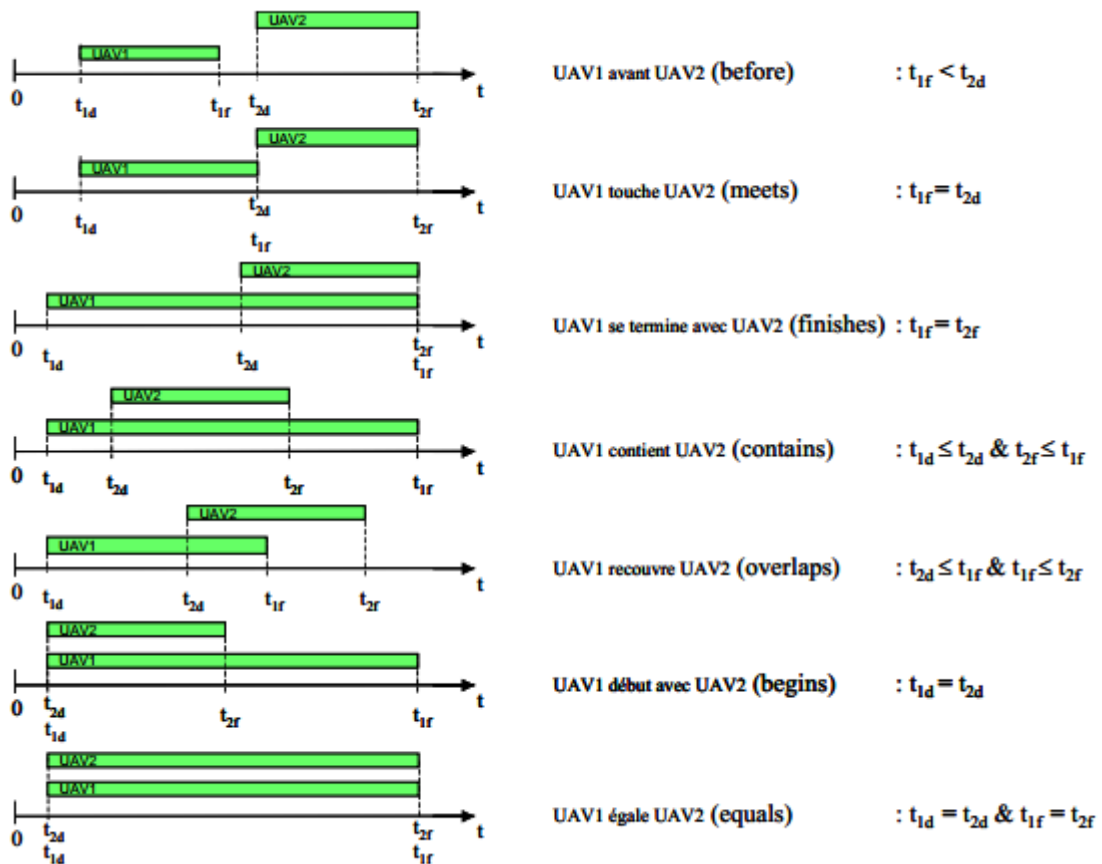


Figure 2. 4 : Les relations temporelles d'Allen

III.2.7. Facette signal

La facette signal regroupe les informations sur le contenu visuel « de bas niveau » telles que : la couleur, la texture, les positions spatiales, etc. Ces informations sont représentées sous formes de descripteurs numériques. Dans plusieurs cas, il est possible d'exploiter ce type d'information pour inférer une description sémantique par l'agrégation d'un certain nombre de ces critères de bas niveau.

La modélisation de la facette signal est inspirée des travaux de M. Belkhatir, [Belkhatir, 05]. La facette signal décrit le contenu visuel du document visioconférence en termes de perception visuelle de l'information visioconférence. Elle permet de spécifier les caractéristiques visuelles de bas niveau de la visioconférence. Formellement, nous dénotons les éléments de cette facette par des descripteurs images (Ids). Ces descripteurs ne sont pas forcément d'ordre symbolique, mais ils peuvent servir à inférer des descriptions sémantiques.

La facette signal est devisée en quatre sous-facettes, [Zwidi *et al.*, 13]:

- la sous-facette couleur pour présenter les caractéristiques couleurs dans le contenu visuel du document ;
- la sous-facette texture permet de décrire les propriétés texture dans le contenu visuel ;
- la sous-facette spatiale pour spécifier les relations spatiales (position relative, direction) des objets images ;
- la sous-facette mouvement pour spécifier les mouvements de la caméra des objets images ainsi que leurs trajectoires.

III.2.7.1. Sous-facette couleur

a. Description

Intégrer les descripteurs signal dans un cadre conceptuel (haut niveau) n'est pas une tâche simple. La première étape consiste à indiquer les données conceptuelles qui correspondent aux descripteurs de bas niveau. Ceci nécessite la mise en place d'un processus de correspondance entre les noms de couleur et les valeurs numériques correspondant à chacune d'elles.

Nous considérons l'existence d'un système formel *Snc*, [Yengui *et al.*, 09], qui distingue un ensemble de catégories couleur avec un *Ccat* (color category). Ces catégories couleur sont distinguées par C_i avec $1 \leq i \leq 11$.

Nous distinguons les 11 couleurs : « C_1 : rouge », « C_2 : blanc », « C_3 : bleu », « C_4 : gris », « C_5 : cyan », « C_6 : vert », « C_7 : jaune », « C_8 : violet », « C_9 : noir », « C_{10} : peau », « C_{11} : orange ».

b. Spécification conceptuelle

Dans le modèle proposé par Belkhatir, l'information de couleur relative à un objet image est décrite par une structure de donnée caractérisant qualitativement et/ou quantitativement sa distribution de couleur ; celui-ci est appelée « méta-concept ». Différentes structures conceptuelles (méta-concepts) sont spécifiées par rapport à différents types de requêtes pouvant être formulées par un utilisateur. Il existe deux types de méta-concepts : booléen et numérique dont chaque type a des sous-types.

L'approche de Belkhatir permet une interrogation riche par l'intermédiaire de trois opérateurs booléen et six opérateurs de quantification appliqués respectivement à des méta-concepts de type booléen ou numérique :

- Un utilisateur a la possibilité d'associer dans une requête des types de concepts de sémantique avec une conjonction booléenne de catégorie couleur.
- Un utilisateur a la possibilité de spécifier des quantifications numériques de catégorie de couleur.

Nous ne reprendrons dans notre modèle que la partie relative aux méta-concepts booléens et aux types de requêtes associées (conjonction, disjonction et négation catégories). Par la suite, « méta-concept » sera synonyme de « méta-concept booléen ». Nous distinguons dans cette

approche des méta-concepts « index » caractérisant la distribution de couleur effectivement présente dans un objet image et des méta-concepts « requête » pour traduire les distributions de couleurs spécifiées au sein des requêtes.

Le méta-concept index pour la couleur présente la distribution des couleurs des objets images par une conjonction de couleur (C_1 à C_{11}) et sont caractérisés par une structure de vecteur C_{ind} avec un nombre d'éléments égal au nombre de catégories de couleur (11 dans notre cas). Les valeurs $C_{ind}[i]$ avec $i \in [1, 11]$ sont des booléens traduisant la présence ou l'absence de la catégorie correspondante C_i dans l'objet image. Les méta-concepts requête sont classés en trois groupes selon le type de relation par lequel ils sont impliqués dans une requête : conjonction, disjonction ou négation. Ils sont caractérisés respectivement par les structures de vecteur C_{et} , C_{ou} et C_{non} avec un nombre d'éléments égal au nombre de catégories de couleur (11). Les valeurs C_{et} , C_{ou} et C_{non} avec $i \in [1, 11]$ sont des booléens traduisant le fait que la catégorie couleur correspondante C_i est élément respectivement des conjonctions, disjonctions et/ou négations des catégories de couleur formulées dans la requête. Ces méta-concepts sont notés comme les méta-concepts index avec en indice à la fin « ET », « OU » ou « NON » indiquant le groupe dont ils font partie. Par exemple :

- Le méta-concept $\langle C_1:1, C_2:1, C_3:1, C_4:0, C_5:0, C_6:0, C_7:0, C_8:0, C_9:0, C_{10}:0, C_{11}:0 \rangle_{ET}$ convient de le représenter aussi par la forme $\langle \text{rouge, blanc, bleu} \rangle_{ET}$.
- Le méta-concept $\langle C_1:0, C_2:0, C_3:1, C_4:0, C_5:0, C_6:0, C_7:0, C_8:0, C_9:1, C_{10}:0, C_{11}:0 \rangle_{OU}$ convient de le représenter aussi par la forme $\langle \text{noir, bleu} \rangle_{OU}$.
- Le méta-concept $\langle C_1:0, C_2:0, C_3:0, C_4:1, C_5:0, C_6:0, C_7:0, C_8:0, C_9:0, C_{10}:0, C_{11}:0 \rangle_{ET}$ convient de le représenter aussi par la forme $\langle \text{gris} \rangle_{NON}$.

Les méta-concepts d'un même groupe peuvent être organisés dans un treillis selon la relation d'ordre partiel « générique > spécifique » selon le fait que si le spécifique est vrai alors le générique l'est aussi. Par exemple :

$\langle \text{rouge, blanc, bleu} \rangle_{ET}$ est un spécifique de $\langle \text{rouge, bleu} \rangle_{ET}$

$\langle \text{noir, bleu} \rangle_{OU}$ est un spécifique de $\langle \text{noir, vert, bleu} \rangle_{OU}$

$\langle \text{gris} \rangle_{NON}$ est un spécifique de $\langle \text{gris, blanc} \rangle_{NON}$

Les trois treillis peuvent être fusionnés en un treillis unique de méta-concepts requête. De même les méta-concepts index peuvent être organisés en un treillis unique (de type ET).

Les méta-concepts index sont mis en relation avec les objets images présents dans le document par la relation « has color ». De même, les méta-concepts requête sont mis en relation avec les objets images spécifiés dans la requête par la même relation « has color ».

Une instance de la sous-facette couleur est représentée par un ensemble de graphes conceptuels chacun comprenant un objet de type I_o lié par la relation conceptuelle « has color » à un objet couleur. Cette représentation est similaire pour le graphe index et le graphe requête. Elle est généralement décrite par le graphe suivant :

$$[I_o] \rightarrow (\text{has color}) \rightarrow [C_{xx}]$$

Exemple :

$$[Io] \rightarrow (\text{has color}) \rightarrow [\langle C_1:1, C_2:0, C_3:0, C_4:0, C_5:0, C_6:0, C_7:0, C_8:0, C_9:0, C_{10}:0, C_{11}:0 \rangle_{ET}]$$

Ou

$$[Io] \rightarrow (\text{has color}) \rightarrow [\langle \text{rouge} \rangle]$$

Notons qu'un objet image dans un document est décrit par une conjonction de couleur. Dans le cas des requêtes, on peut aussi avoir une disjonction de couleur ($\langle \text{col} \rangle_{OU}$). Par exemple : « rechercher des objets images ayant une couleur rouge ou bleu ».

La représentation des cas de conjonction et disjonction des couleurs par des graphes conceptuels et la suivante :

$$[Io] \rightarrow (\text{has color}) \rightarrow [\langle \text{col} \rangle_{ET}]$$

$$[Io] \rightarrow (\text{has color}) \rightarrow [\langle \text{col} \rangle_{OU}]$$

Avec $\langle \text{col} \rangle_{ET}$ et $\langle \text{col} \rangle_{OU}$ représentant respectivement une combinaison de 11 valeurs booléennes représentant les concepts couleurs.

III.2.7.2.Sous-facette texture

a. Description

Notre représentation symbolique de la texture se base sur les travaux de recherche en nommage et catégorisation de textures proposés par [Bhushan *et al.*, 97]. Nous considérons les textures suivantes comme représentation de chacune de ces catégories : « T₁ : bosselé », « T₂ : craquelé », « T₃ : désordonné », « T₄ : entrelacé », « T₅ : rayé », « T₆ : marbré », « T₇ : réiforme », « T₈ : sali », « T₉ : tacheté », « T₁₀ : uniforme » et « T₁₁ : en vague ». Ces textures sont désignées par T_i avec $1 \leq i \leq 11$.

b. Spécification conceptuelle

La forme de la spécification conceptuelle est exactement la même pour la texture que pour la couleur. Les 11 textures de bases jouent le même rôle que les 11 couleurs de base et elles peuvent être combinées de la même façon pour produire des treillis de méta-concepts texture index et de méta-concepts texture requête. La relation « has color » a un équivalent noté « has texture ». Les mêmes notations concises peuvent être utilisées.

Une instance de la sous-facette texture est représentée par un ensemble de graphes conceptuels chacun comprenant un objet de type Io lié par la relation conceptuelle « has texture » à un méta-concept texture. Cette représentation est similaire pour le graphe index et le graphe requête. Elle est généralement décrite par le graphe suivant :

$$[Io] \rightarrow (\text{has texture}) \rightarrow [T_{xx}]$$

Si par exemple l'objet Io a une texture rayée, une représentation dans le formalisme des graphes conceptuels sera comme suite :

$$[Io] \rightarrow (\text{has texture}) \rightarrow [\langle T_1:0, T_2:0, T_3:0, T_4:0, T_5:1, T_6:0, T_7:0, T_8:0, T_9:0, T_{10}:0, T_{11}:0 \rangle_{ET}]$$

Ou

$$[Io] \rightarrow (\text{has texture}) \rightarrow [\langle \text{rayé} \rangle]$$

Notons qu'un objet image dans un document peut être décrit par une conjonction de textures. Dans le cas des requêtes, on peut aussi avoir une disjonction de textures ($\langle \text{tex} \rangle_{OU}$). Par exemple : « rechercher les objets images ayant une texture Sali ou désordonnée ».

La représentation des cas de conjonction et disjonction des couleurs par des graphes conceptuels est la suivante :

$$[Io] \rightarrow (\text{has texture}) \rightarrow [(\langle \text{tex} \rangle_{ET})]$$

$$[Io] \rightarrow (\text{has texture}) \rightarrow [(\langle \text{tex} \rangle_{OU})]$$

Avec $\langle \text{tex} \rangle_{ET}$ et $\langle \text{tex} \rangle_{OU}$ représentant respectivement une combinaison de 11 valeurs booléennes représentant les concepts texture.

III.2.7.3. La sous-facette spatiale

a. Description

La facette spatiale permet de spécifier une description spatiale d'un objet par rapport aux autres du point de vue de la perception visuelle. D'autre part, déterminer la position géographique d'un objet vidéo (déterminer le lieu où une telle scène se déroule).

La description spatiale d'une séquence vidéo se traduit selon deux formes. La première au niveau topologique permettant de situer les objets les uns par rapport aux autres. Ceci résulte de la perception visuelle du contenu. Il est souvent qualifié de statique du fait qu'il s'applique aux images fixes. On considère 10 relations qui sont : « à gauche de », « à droite de », « au-dessous de », « au-dessus de », « devant », « intérieur de », « touche », « déconnecté de », « proche de », « loin de ».

Afin de modéliser des données spatiales, nous considérons d'abord un sous-ensemble des relations topologiques. Considérons deux objets images (Io1 et Io2), ces relations sont :

(S₁=C, « devant ») : « Io1 est devant Io2 »

(S₂=P, « est intérieur de ») : « Io1 est intérieur de Io2 »

(S₃=T, « touche ») : « Io1 touche Io2 »

(S₄=D, « déconnecté de ») : « Io1 est déconnecté de Io2 »

Des relations directionnelles qui sont invariables par rapport aux transformations géométriques de base :

(S₅, à droite de) « Io1 est à droite de Io2 »

(S₆, à gauche de) « Io1 est à gauche de Io2 »

(S₇, au-dessus de) « Io1 est au-dessus de Io2 »

(S₈, au-dessous de) « Io1 est au-dessous de Io2 »

Deux autres relations basées sur les distances entre les objets d'image qui sont :

(S₉, proche de) « Io1 est proche de Io2 »

(S₁₀, loin de) « Io1 est loin de Io2 »

Une instance de la facette spatiale est représentée par un ensemble de GCs, chacun contenant deux types d'objets images liés par les relations spatiales précédemment définies.

b. Spécification conceptuelle

Les objets images sont représentés par des concepts de type Io. Les relations spatiales sont représentées dans un treillis de relations SpC (cf. **Figure 2.5**). Chaque pair d'Io est interconnecté par l'une des relations du SpC.

Une instance de la sous-facette spatiale est représentée par un ensemble de graphes conceptuels chacun comprenant un concept de type Io lié par la relation spatiale. Cette représentation est généralement décrite par le graphe suivant :

$$[Io] \rightarrow (SpC) \rightarrow [Io2]$$

Considérant Io1 et Io2 deux objets images représentant respectivement les objets *Dr Joël* et *le patient* de l'exemple suivant : « les segments visioconférence montrant un expert explique une opération chirurgicale au cours ». Une représentation avec le formalisme des graphes conceptuels en exploitant une description spatiale est décrite par le graphe ci-dessous :

$$[Io1] \rightarrow (\text{devant}) \rightarrow [Io2]$$

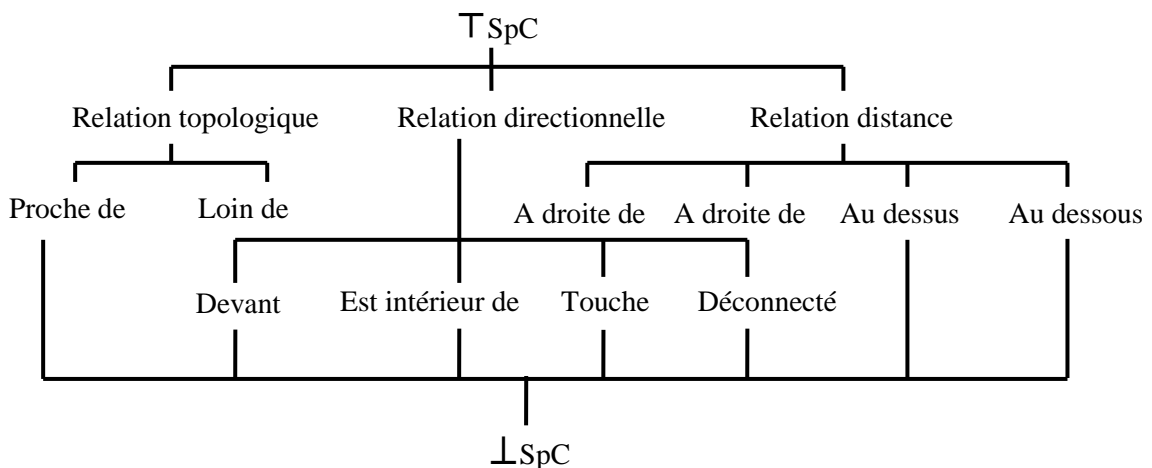


Figure 2. 5: Classification de relations conceptuelles

III.2.7.4. La sous-facette mouvement

a. Description

Plusieurs travaux ont été proposés pour le développement des algorithmes et des techniques pour l'estimation des mouvements de caméra et les trajectoires des objets dans le document vidéo. Ces travaux restent à un niveau signal et peu de propositions ont été faites pour la caractérisation des mouvements. Dans le cas de la visioconférence, nous distinguons deux types de mouvements. Le premier concerne la description des mouvements de la caméra (considérés au niveau image) et le deuxième décrit les mouvements des objets, [Wang et al., 11].

Dans le cas où la caméra n'est pas fixe, un mouvement d'objet peut être « absolu » s'il est estimé par rapport un « fond » ou bien « relatif » s'il est estimé par rapport à une image. Dans le cas inverse (caméra fixe) les notions de « mouvement objet relatif » et « mouvement objet absolu » sont fusionnées.

Un mouvement général peut être décomposé en mouvements élémentaires. Ces mouvements élémentaires peuvent être : une « translation », une « rotation », une « homothétie » ou bien « irrégulier » (non linéaire). Excepté le mouvement « irrégulier » (qui permet uniquement de décrire le mouvement d'objets non rigides), les mouvements élémentaires peuvent être appliqués à la description des mouvements de la caméra et aussi des objets.

Nous distinguons huit mouvements élémentaires qui forment la base de notre cadre pour la caractérisation symbolique de l'aspect « mouvement » dans la visioconférence : « M_1 = translation horizontale », « M_2 = translation verticale », « M_3 = translation en profondeur », « M_4 = rotation autour de l'axe horizontal », « M_5 = rotation autour de l'axe vertical », « M_6 = rotation autour de l'axe de profondeur », « M_7 = homothétie » et « M_8 = transformation irrégulière ».

Ces huit mouvements élémentaires peuvent être présents ou absents indépendamment les uns des autres et toutes les combinaisons de ceux-ci sont possibles. Les sept premiers correspondent à des degrés de liberté soit de la caméra soit de l'objet (bien que le dernier, « gonflement / rétrécissement », soit exceptionnel dans ce cas).

Le huitième n'existe que dans le cas des objets et traduit une déformation de ceux-ci qui ne se réduit pas aux sept premiers mouvements élémentaires (mouvement de foule, vagues, personne qui marche ou se penche...). Nous pouvons toutefois le considérer comme indépendant des sept premiers dans la mesure où (en théorie au moins) nous pouvons extraire un mouvement de translation, de rotation ou de grossissement même dans le cas de mouvements irréguliers (déplacement du centre de gravité pour la translation par exemple). Nous pouvons toujours trouver une « meilleure translation », une « meilleure rotation » et une « meilleure homothétie », celles-ci permettent de déterminer l'absence ou la présence (et même l'orientation) des sept premiers mouvements élémentaires.

Les sept premiers mouvements « mouvement élémentaires » correspondent à la nomenclature utilisée pour décrire les mouvements de la caméra (cf. **Figure 2.6**). Elles correspondent respectivement à : « track », « boom », « dolly », « tilt », « pan », « roll » et « zoom ». Une description par les mouvements est associée à l'image pour les mouvements de caméra et aux Io pour les mouvements des objets. Nous gardons la notation Io qui fait référence à la notion d'objet image dans le plan.

b. Spécification Conceptuelle

La forme de la spécification conceptuelle est similaire à celle utilisée pour la texture et la couleur. La principale différence concerne le fait qu'une caractéristique élémentaire de mouvement peut être non seulement présente ou absente mais si elle est présente, elle peut être dans une direction ou dans une autre (vers la gauche ou la droite par exemple). Les 8

mouvements de base jouent le même rôle que les 11 couleurs ou les 11 textures de base et elles peuvent être combinées de la même façon pour produire des treillis de méta-concepts texture index et de méta-concepts texture requête.

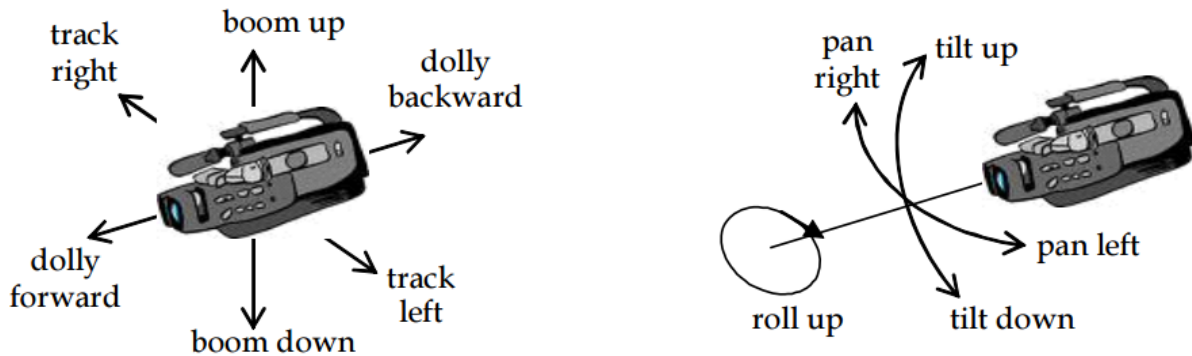


Figure 2. 6: Les mouvements de caméra

Les relations « has color » et « has texture » ont un équivalent noté « has motion ». Il y a par contre dans le cas du mouvement une extension de l'ensemble des valeurs possibles pour les composantes des vecteurs correspondants aux mouvements élémentaires. En plus du « 0 » et du « 1 » qui ont la même signification (présence ou absence du mouvement élémentaire) nous introduisons les valeurs « -1 » et « +1 » qui spécifient à la fois la présence du mouvement élémentaire et son orientation. Il convient de noter que nous utilisons ici des symboles qui ne s'identifient pas à des nombres entiers. En particulier « +1 » est différent de « 1 ». Des notations concises peuvent également être utilisées. Le **tableau 2.1.** indique la correspondance entre les symboles « -1 » et « +1 » et l'orientation correspondante pour les sept premiers mouvements élémentaires.

Les méta-concepts mouvement peuvent encore être organisés en treillis car une relation d'ordre partiel entre eux peut également être définie. Celle-ci généralise celle qui est utilisée dans le cas des méta-concepts de couleur ou de texture. En ce qui concerne la relation générique – spécifique, les valeurs de composante « -1 » ou « +1 » induisent une relation de spécificité par rapport à un méta-concept ayant la valeur « 1 » et identique par ailleurs. Ces relations d'ordre partiel permettent encore de déterminer si un méta-concept présent dans une requête peut être projeté sur (s'il généralise) un méta-concept présent dans l'index d'un segment de visioconférence.

Une instance de la sous-facette mouvement est représentée par un ensemble de graphes conceptuels chacun comprenant un concept de type Io lié par la relation conceptuelle « has motion » à un méta-concept mouvement. Cette représentation est similaire pour le graphe index et le graphe requête. Elle est généralement décrite par le graphe suivant :

$$[Io] \rightarrow (\text{has motion}) \rightarrow [M_{xx}] \text{ (pour les objets image)}$$

ou

[VS] → (has motion) → [M_{xx}] (pour la caméra, VS : Video Segment)

Tableau 2. 1 : Description des différents types de mouvement

Type de mouvement	Valeur	signification
Translation horizontale	-1	translation vers la gauche
	+1	translation vers la droite
Translation vertical	-1	translation vers le bas
	+1	translation vers le haut
Translation en profondeur	-1	translation en arrière
	+1	translation en avant
Rotation autour d'un axe horizontal	-1	rotation vers le bas
	+1	rotation vers le haut
Rotation autour d'un axe vertical	-1	rotation vers la gauche
	+1	rotation vers la droite
Rotation autour d'un axe en profondeur	-1	rotation sens inverse d'aiguille d'une montre
	+1	rotation au sens d'aiguille d'une montre
Homothétie	-1	réduire
	+1	agrandir

Si par exemple l'objet Io a un mouvement de translation à gauche, une représentation dans le formalisme des graphes conceptuels sera comme suite :

[Io] → (has motion) → [<M₁: -1, M₂: 0, M₃: 0, M₄: 0, M₅: 0, M₆: 0, M₇: 0, M₈: 0>]

ou

[Io] → (has motion) → [<translation à gauche>]

Notons qu'un objet image dans un document peut être décrit par une conjonction de mouvements élémentaires. Par exemple, un objet image peut se déplacer à la fois vers la gauche et vers le haut tout en tournant dans le sens des aiguilles d'une montre. Dans le cas des requêtes, nous pouvons aussi avoir une disjonction de mouvements (<mvt>_{OU}). Par exemple, « rechercher les objets images se déplaçant vers la droite ou vers le bas ».

La représentation des cas de conjonction et disjonction des couleurs par des graphes conceptuels est la suivante :

[Io] → (has motion) → [<mvt>_{ET}]

[Io] → (has motion) → [<mvt>_{OU}]

Avec <mvt>_{ET} et <mvt>_{OU} représentant respectivement une combinaison de 8 valeurs booléennes représentant les concepts texture. Par exemple :

Une requête spécifiant un Io ayant un mouvement de translation horizontale (gauche ou droite mais non nulle) et vers le bas contient le sous-graphe :

$$[Io] \rightarrow (\text{has motion}) \rightarrow [<M_1:1, M_2:-1, M_3: 0, M_4:0, M_5: 0, M_6: 0, M_7:0, M_8:0>_{ET}]$$

Une représentation de la requête « rechercher un mouvement de caméra panoramique à droite ou zoom out » contient le graphe suivant :

$$[VS] \rightarrow (\text{has motion}) \rightarrow [<M_1:+1, M_2:0, M_3: 0, M_4:0, M_5: 0, M_6: 0, M_7:0, M_8:-1>_{OU}]$$

Notons que notre description des objets en mouvement se limite à la description de mouvement apparent des objets « rigides ». Une indexation des objets « déformable » plus précise que « mouvement irrégulier » nécessite une analyse plus fine qui consiste à décomposer l'objet en des sous-objets « non déformables ». Une possibilité d'extension de notre proposition est donc envisageable par la prise en compte des objets « déformables » pour la modélisation de la sous-facette mouvement. À partir de la modélisation des mouvements de ces objets, nous pouvons inférer des descriptions sémantiques associées telles que marcher, se pencher, ...

Dans le cas des objets « non rigides » où le mouvement peut être segmenté en des multiples phases durant un même plan, une micro segmentation du plan est nécessaire afin d'extraire des mouvements élémentaires, [Dumont *et al.*, 12] et [Khoury *et al.*, 14].

Un autre cas important pour la modélisation du mouvement et que nous n'avons pas pris en compte dans notre modèle consiste à décrire le mouvement réel dans une scène entre deux (ou plus) objets par d'autres informations symboliques telles que par exemple « se rapprocher », « s'éloigner », etc.

III.3. Modélisation du contenu audio

Le contenu audio constitue une base pour un grand nombre d'approches présentées dans la littérature. Ces approches se distinguent par la spécificité du cadre applicatif généralement restreint à des applications données.

III.3.1. Spécification conceptuelle et description du contenu audio

La description du contenu audio a pour objectif de permettre aux utilisateurs de formuler des requêtes sous une forme simple (en langage naturel). Dans ce contexte, l'utilisation de simples mots clés ne permet pas d'obtenir la meilleure précision possible. Nous choisissons donc de mettre en avant une description conceptuelle et relationnelle du contenu. Un terme qui apparaît dans le graphe fait référence à une description symbolique du contenu. Nous distinguons des cas spécifiques c.à.d des graphes destinés à représenter un seul type de flux et ceci par l'existence des relations conceptuelles implicites ou explicites.

Dans la **figure 2.7.**, nous donnons une représentation abstraite du contenu audio sous forme graphique en utilisant le formalisme des GCs.

Une description conceptuelle n'est pas forcément la conséquence d'une interprétation par un opérateur humain de ce qu'il voit ou bien de ce qu'il entend. En effet, il existe des descriptions conceptuelles qui résultent de l'analyse du contenu signal pour la détection des actions spécifiques telles que par exemple une action d'applaudissement.

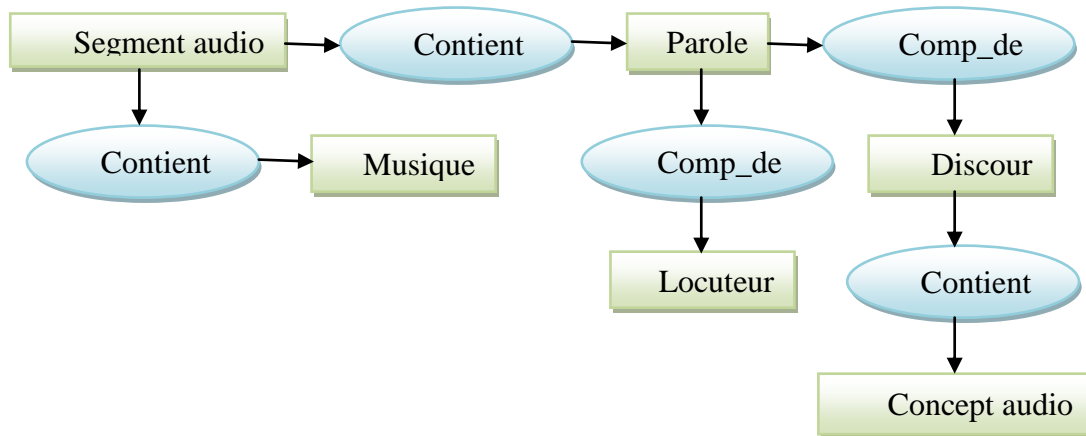


Figure 2. 7: Description générique du contenu audio

III.3.2. Segmentation audio et structure du document

L'objectif de la segmentation basée sur le contenu audio est d'obtenir des segments cohérents. Chaque segment possède un contenu spécifique : musique, bruit ou parole, [Vincent, 12] et [Parvaix, 10]. Pour la parole, un autre type de segmentation est possible comme la segmentation du document suivant le tour de parole (changement de locuteur où chaque segment devra contenir la parole prononcé par une personne), [Charhad, 05] et [Bendris, 11].

Le contenu audio permet de décrire toute information : soit de compléter et d'enrichir les descriptions issues du contenu visuel, soit d'apporter une interprétation indépendante du contenu du segment là où le contenu visuel ne le permet pas.

Un objet audio peut être n'importe quel élément sémantique du contenu audio, souvent des termes extraits automatiquement à partir de l'audio tel que par exemple l'identité d'une personne, le lieu, etc. L'information sémantique contient parmi d'autres deux éléments d'informations particuliers. Le premier élément consiste à spécifier qui parle dans chaque segment visioconférence (spécifié par la relation parle). Le deuxième élément concerne de l'information dont on parle (spécifié par la relation parle de).

Il existe donc deux relations conceptuelles audio (« parle » et « parle de »). La distinction entre ces deux relations conceptuelles, bien qu'elles puissent exprimer la même signification, permet de tenir compte de la spécificité des requêtes.

III.3.3 Modélisation de la sous-facette audio

Les objets audio seront présentés dans la suite par les concepts Ao. Chaque segment visioconférence est représenté par un ensemble de graphes conceptuels comme dans le cas de la modélisation du contenu visuel. Chaque graphe contient des Ao interconnectés par des relations conceptuelles qui sera présenté comme [Ao1] → (parle de) → [Ao2] ou bien [Ao1] → (parle). Dans le cas d'une visioconférence télé-chirurgie, par exemple, nous trouvons trois types de participants : l'équipe qui opère un patient dans une salle d'opération, un expert connecté par visioconférence et qui reçoit les images de l'opération en cours et fournit des

conseils en ligne et un public, composé de chirurgiens en formation continu qui assiste à l'opération en direct et peut poser des questions au chirurgien responsable de l'opération aussi bien qu'à l'expert. Le public peut entendre l'intervention de l'expert sans le voir vu que la caméra sera fixée sur l'opération chirurgicale.

L'information sémantique dans le contenu audio contient deux éléments d'information particuliers : le premier consiste à spécifier qui parle dans chaque segment présenté par la relation « parle », le deuxième représente l'information dont on parle présenté par la relation « parle de ». La distinction entre ces deux relations conceptuelles bien qu'elle puisse exprimer la même signification, permet de tenir compte de la spécificité des requêtes. Si nous faisons référence à la télé-chirurgie indiquée au-dessus, l'expert fournit des conseils en ligne concernant l'opération en cours. Parmi les concepts existants dans cette visioconférence, nous trouvons le concept « Expert » et le concept « Opération en cours ». Ces deux concepts sont reliés par la relation « parle de ». La distinction « L'expert parle de l'opération en cours » peut être formulée par les deux graphes tels que :

[Expert] → (parle)

et

[Expert] → (parle de) → [Opération en cours]

Si nous considérons par exemple deux objets audio (Ao1 et Ao2), ces relations sont (a1= Ao1, parle) avec Ao1 est dans la classe de concept personne, (a2= Ao1, parle de, Ao2) qui se traduit par « Ao1 parle de Ao2 ». Ao2 peut être n'importe autre concept cité dans le discours. Nous distinguons trois catégories de concept sémantique audio (Asc) : (Asc1= personne, Asc2 = opération médical, Asc3 = lieu). Ces concepts sont classés dans des treillis comme montré dans la **figure 2.8**.

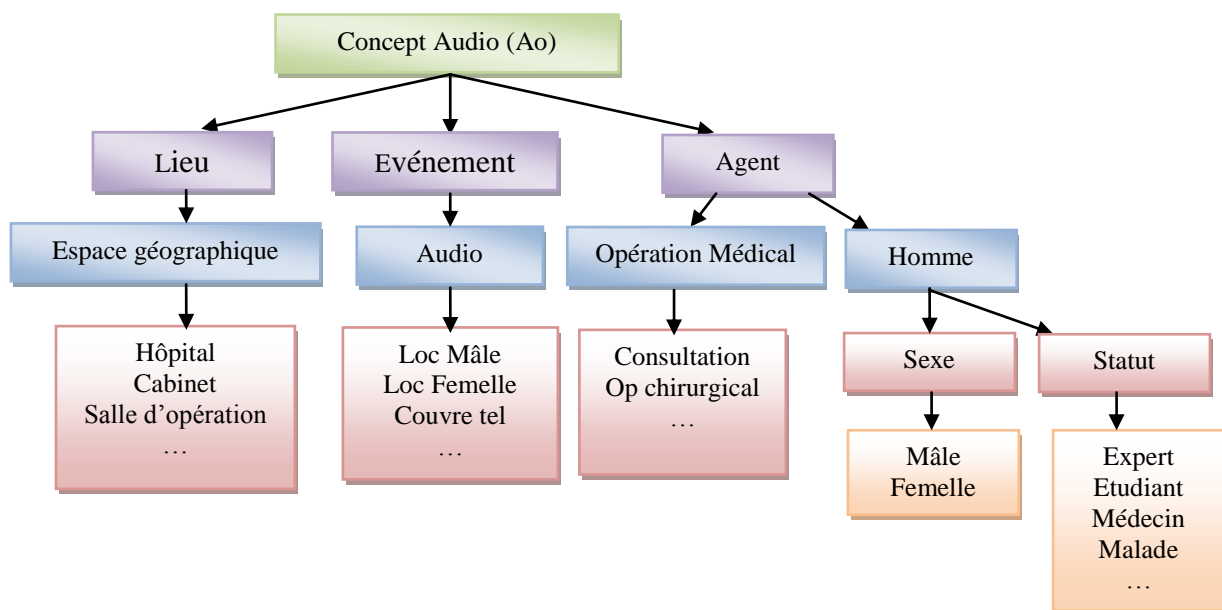


Figure 2. 8: Treillis des concepts audio

IV. Modélisation sémantique

IV.1. Modélisation du contenu visuel : Facette sémantique

La facette sémantique permet de modéliser l'ensemble des descriptions de haut-niveau dans un document visioconférence. Nous distinguons trois sous-facettes dont chacune décrit un type de média spécifique (le visuel, l'audio et le texte).

Cette facette permet de décrire l'information visuelle contenue dans le document visioconférence. La sémantique décrite dans cette facette est portée par des objets images (Io) qui correspondent à des concepts.

IV.1.1. Modèle de la sous- facette visuelle

Si les images fixes sont décrites uniquement par les aspects physiques de l'image tels que la couleur, la texture, etc.; les images animées sont plus riches au niveau sémantique. En effet, la description du mouvement des objets et le suivi de leurs trajectoires ainsi que la description de mouvement de caméra (panoramique, zoom,...) sont des éléments non négligeables pour comprendre le contenu de la visioconférence, [Wang *et al.*, 11].

L'intégration de l'information au niveau signal et au niveau conceptuel est cruciale puisqu'elle enrichit la structure d'indexation et permet de combiner les aspects signal et sémantique dans le module de recherche.

Les objets image sont représentés par des concepts Io et des concepts sémantiques visuels qui sont organisés dans un treillis ordonné par de relation de type spécifique/générique. Un exemple de la sous-facette visuelle est représenté par un ensemble de graphes conceptuels, chacun contenant un type de Io lié à un Visual Semantic Concept (VSC) par une relation conceptuelle (apparaît). Nous nous basons dans ce sur la perception visuelle humaine afin d'interpréter ces informations. Ce type de modélisation permet en fait de répondre à des requêtes spécifiques au contenu visuel.

Le graphe de base pour la génération des représentations liées à cette facette est :

$$[Io] \rightarrow (\text{apparaît}) \rightarrow [VSC].$$

IV.1.2. Extraction de concepts visuels

Les VSC résultent de l'interprétation du contenu visuel des objets images. Ceux-ci sont extraits pour établir une liste de VSC qui peut être utilisée pour indexer le plan vidéo représenté par une image. L'objectif principal est de fournir des indexations collaboratives qui vont servir à une base d'apprentissage pour l'extraction des traits visuels (*cf.* **Figure 2.9.**).

IV.2. Modélisation du contenu Audio : extraction des termes

Le processus d'extraction consiste à détecter les termes dans un contexte documentaire. Un contexte documentaire est défini comme une unité textuelle à l'intérieur d'un document XML. Il peut représenter une phrase, un paragraphe ou un élément logique de la structure logique (les nœuds texte dans les documents XML). D'abord, pour chaque concept (ou un des synonymes de ce concept) nous chercherons sa présence dans le document traité. Ensuite, nous calculerons le nombre d'occurrence de chaque concept dans le document

qui est le cumule de tous les concepts trouvés ainsi que leur synonymes. La fréquence d'apparition de chaque concept dans le document est égal au nombre d'occurrence de chaque concept dans le document sur le nombre total des concepts du document, [Baccari *et al.*, 12].

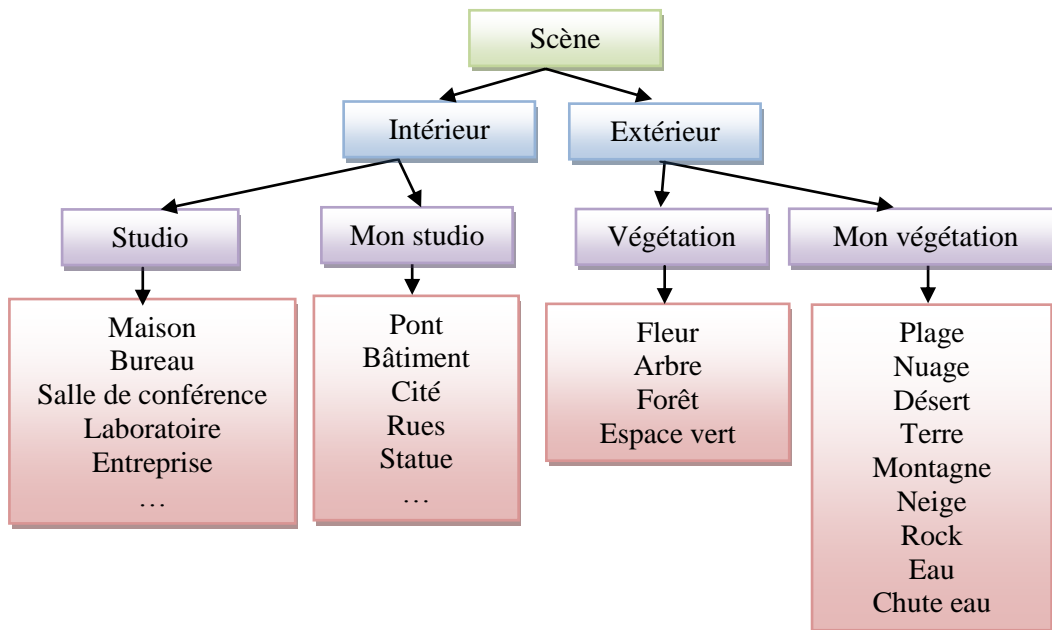


Figure 2. 9 : Treillis de concepts visuels

L'objectif de cette étape est d'extraire tous les termes du document susceptibles de représenter des concepts dans l'ontologie. Ces termes correspondent à différentes entrées (ou nœuds) dans l'ontologie. À cet effet, nous utilisons une technique qui consiste à projeter l'ontologie sur le document. Cela est fait par un parcours de l'ontologie en utilisant un parseur développé à cet effet afin d'identifier les concepts de l'ontologie qui ocurrent comme des termes dans le document, [Baccari *et al.*, 12].

Nous fixons pour cela deux objectifs dont le premier consiste à extraire les termes simples à partir des documents XML. Et à partir de ces termes, nous dégageons les termes composés (*cf.* **Figure 2.10.**).

IV.2.1. Transcription de la parole

Le traitement automatique de la parole est l'une des technologies déterminantes pour le développement d'interfaces Homme-Machine avancées. Dans une visioconférence, l'analyse de la partie audio permet une caractérisation qui peut, dans la majorité des cas, donner des résultats très précis sur le contenu de cette visioconférence.

Nous essayons de ne retenir de la source que ce qui peut le mieux la caractériser et la rendre unique. En fait, nous allons extraire une signature caractérisant d'une manière unique un extrait quelconque d'une bande audio. Cette étude va être analogue à celle d'une signature image, construite pour la même raison. Enfin, nous combinons les deux signatures pour

qu'elles se complètent et ainsi obtenir une signature globale très forte, pour un document visioconférence.

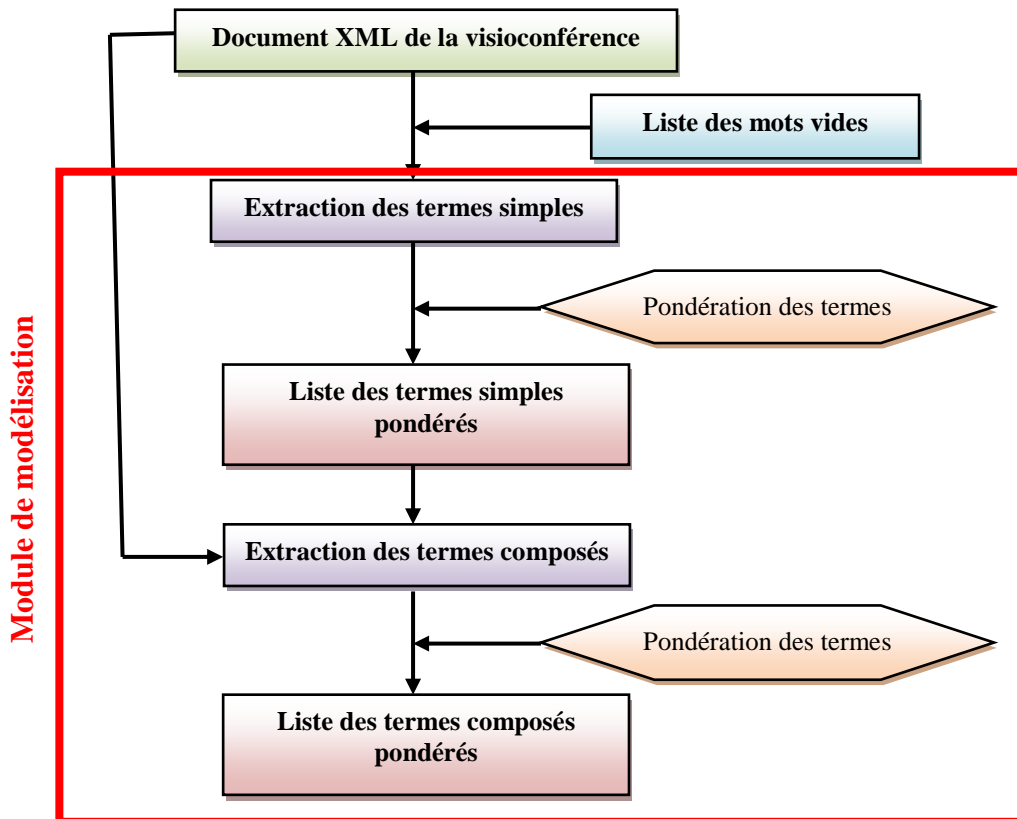


Figure 2. 10 : Module de modélisation (Analyse)

Généralement, dans un système qui indexe du contenu audio, la première étape consiste à extraire les caractéristiques de bas niveau sur lesquelles vont se baser les autres algorithmes d'indexation. C'est alors la segmentation et la catégorisation du son qui sont privilégiées. Les classes les plus utilisées sont : les silences, les dialogues (paroles), les bruits et la musique. Notons que la segmentation du contenu audio ne correspond pas forcément à celle du contenu visuel d'un point de vue de la durée et des bornes des intervalles.

Les informations portées par la parole sont très utiles au module d'indexation des visioconférences. En effet, une partie des index est directement extraite des transcriptions automatiques analysées par des méthodes du domaine du TALN. Ces méthodes permettent, par exemple de détecter des mots-clés caractéristiques, des champs lexicaux exprimant une thématique ou des entités nommées (noms propres, lieux, dates...), [Bigot, 11].

Pour la transcription des discours, ce n'est qu'une transcription orthographique consistant en une reconnaissance de la séquence de mots prononcée dans un flux de parole. Les systèmes de transcription traitent de la parole continue et reconnaissent un vocabulaire étendu. Ils sont

indépendants des locuteurs, [Bougares, 12]. Après la transcription des discours, nous nous aidons de l'indexation des documents textuels pour indexer ce discours transcrit.

Cette transcription de la parole permet d'extraire des informations sur les thématiques traitées. En fait, les descripteurs, à travers l'intonation et le débit de parole peuvent être révélateurs d'intentions, d'opinions des intervenants. Notons que l'analyse textuelle des sous-titres peut également permettre d'obtenir des informations sur les intervenants.

Nous proposons les étapes présentées par la **figure 2.11** pour achever une indexation du contenu auditif.

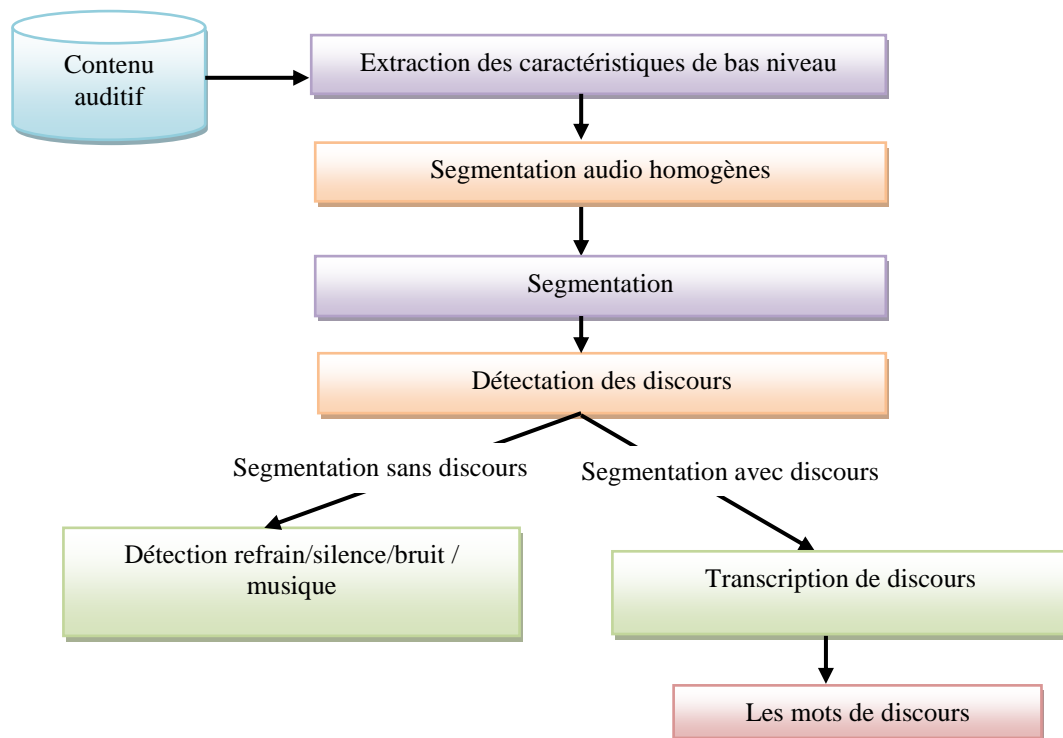


Figure 2. 11: Etapes de l'indexation audio

La transcription automatique est effectuée en deux passes, dont la première sert à générer rapidement une séquence de mots approximative. Celle-ci permet une adaptation en aveugle des modèles acoustiques. Les modèles ainsi adaptés sont utilisés en deuxième passe.

La modélisation du contenu audio dépend uniquement du texte transcrit. Toutefois, nous aurons forcément lors de la transcription des erreurs résultantes de la reconnaissance automatique qui peut influencer la pertinence des réponses. Dans une visioconférence, les données orales illustrent différents styles de parole (préparée et spontanée). De nombreuses erreurs de **R**econnaissance **A**utomatique de la **P**arole (RAP) sont dues à des mots fréquents homophones ayant même prononciation mais différentes écritures et différents sens. Pour ces mots, la solution correcte dépend du modèle de langage. Une **C**lassification **A**utomatique (CA) a été effectuée pour discriminer deux paires homophones ('et/est', 'à/a' ...) par des propriétés acoustiques et prosodiques.

Dans le texte résultat de cette transcription, l'information apparaît comme un message à transmettre le plus fidèlement possible, en cherchant à limiter la déperdition. Elle est considérée comme un signal mais surtout un stimulus qui conduit à une réponse chez une personne. Le sens se « construit » par des opérations d'interprétation (en fonction des connaissances antérieures, du type de cheminement, du locuteur). D'où l'importance de s'attacher à caractériser au mieux les informations dont nous disposons pour donner à l'utilisateur le maximum de clarté sur l'information. L'étroite interdépendance entre le locuteur, son action et l'information est constitutive du rapport à l'information. L'individu a une certaine relation à l'information, et une intention par rapport à une action (Il veut pouvoir se cultiver, discuter, négocier...), [Guyot, 11].

Les transcriptions automatiques fournies ne sont cependant pas parfaites, et sont accompagnées toujours avec certains problèmes sans solutions satisfaisantes. En effet, les systèmes arrivent à fournir des transcriptions avec un haut niveau de précision lorsque la parole est préparée (très proche d'une parole lue), mais voient leurs performances chuter fortement lorsque la parole est dite spontanée, apparaissant lors de conversations non préparées (dialogues, débats. . .).

Les zones de conversation, sont intrinsèquement porteuses d'une grande quantité d'informations sur la nature du document. Nous proposons dans ce sens de caractériser les conversations afin de faire ressortir leurs particularités. La caractérisation des zones de conversation doit être indépendante de la transcription. Ceci afin de s'affranchir des erreurs inhérentes aux méthodes automatiques sur ce type de contenu, [Guyot, 11].

Par conclusion, ces dernières années, il y a des travaux qui sont faits pour une indexation pertinente des éventuelles conversations et d'autres pour trouver des solutions concernant les mots qui se prononcent de la même façon et d'autres qui distinguent les paroles spontanées des paroles préparées. À titre d'exemple, nous citons le travail de Guyot « *Introduction à l'ingénierie documentaire et aux sciences de l'information* », [Guyot, 11], le travail de Bigot « *Recherche du rôle des intervenants et de leurs interactions pour la structuration de documents audiovisuels* », [Bigot, 11] et le travail de Dufour « *Transcription automatique de la parole spontanée* », [Dufour, 10].

Or, ces travaux sont faits séparément les uns aux autres et les visioconférences présentent la spécificité de contenir à la fois des paroles spontanées et d'autres préparées, des conversations lors d'une éventuelle discussion ayant lieu après une séquence d'une visioconférence portant par exemple sur une opération chirurgicale présentant l'avantage de contenir des zones de parole superposées. Nous allons profiter des résultats achevés dans chacun de ces travaux.

Pour limiter l'imperfection des transcriptions automatiques fournies, nous essayons d'améliorer les performances des systèmes de RAP. Nous ajoutons au niveau du dictionnaire des prononciations spécifiques. Nous corrigeons certains mots homophones, des erreurs spécifiques en post-traitement des transcriptions.

Faut-il signaler qu'une conversation dans une visioconférence peut contenir des pauses regroupant des morphèmes tels que "euh", "hum". La spécification de cette fonction morphèmes spécifiques, reste difficile à définir. Le morphème "euh" est alors catégorisé en tant qu'hésitation, mais pour les autres morphèmes ("ben",...), une catégorisation reste plus délicate.

En outre, elle peut, aussi, contenir des pauses non-sonores marquant souvent, la rupture au niveau d'une idée. Ces pauses permettent de structurer le discours, [Bazillon, 11].

Or, La description du contenu audio a pour objectif de permettre aux utilisateurs de formuler des requêtes sous une forme simple c'est-à-dire en langage naturel. Nous choisissons, alors, une description conceptuelle et relationnelle du contenu pour obtenir la meilleure précision possible qui ne peut pas être achevée lors d'une utilisation de simples mots clés.

Nous donnons dans la **figure 2.12.** une représentation abstraite du contenu audio sous forme graphique en utilisant le formalisme des graphes conceptuels.

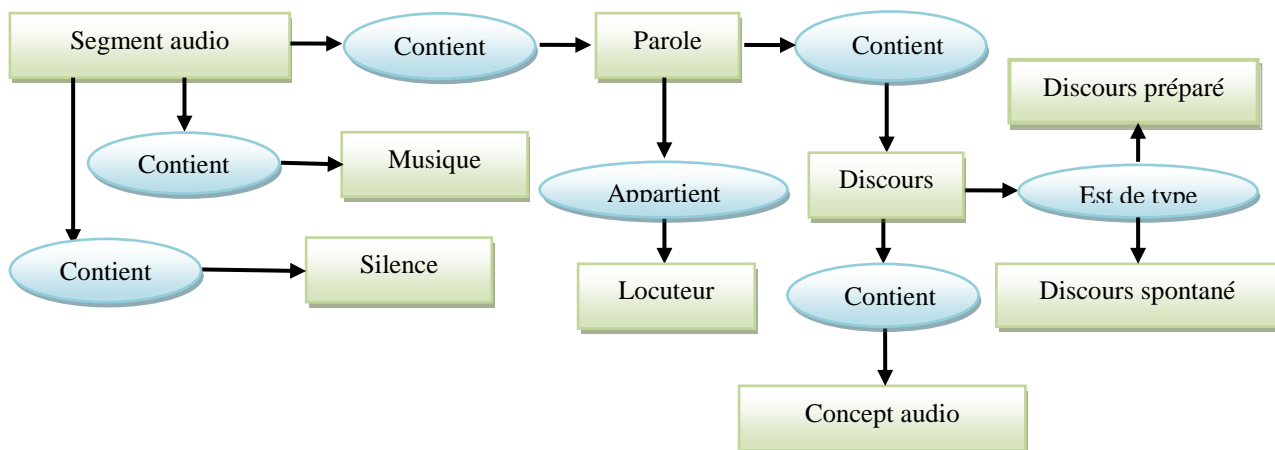


Figure 2. 12: description générique du contenu audio

IV.2.2. Extraction des termes simples

IV.2.2.1. Extraction des mots vides candidats

Un mot vide candidat est un mot qui est tellement commun qu'il est inutile de l'indexer ou de l'utiliser. Dans cette étape, nous affectons à chaque mot une catégorie : vide ou plein. Les mots vides (ou stop word en anglais) sont des mots qui sont communs à tous les textes dans une même langue. Ils ont une utilité fonctionnelle. En français, les mots vides évidents pourraient être « le », « la », « de », « du », « ce », « ça », etc. Dans un contexte monolingue où tous les documents du corpus sont dans une même langue, les mots vides sont principalement des mots caractéristiques de cette langue tels que les pronoms, les prépositions, les articles, etc. Dans ce contexte, les mots vides sont dits encore mots grammaticaux. Alors, il est inutile de les indexer ou de les utiliser dans un module de

recherche d'information. Dans un texte, un mot vide est un mot non significatif contrairement à un mot plein, [Zwidi *et al.*, 13].

IV.2.2.2. Extraction des termes simples par élimination des mots vides

Afin d'extraire les termes simples, nous procédons par élimination des mots vides (les conjonctions, les prépositions, certains adverbes, etc.). L'utilisateur peut modifier la liste des mots vides selon son choix. L'ensemble des mots du corpus est constitué de deux sous-ensembles : un sous-ensemble de mots vides et un sous-ensemble de mots pleins comme termes simples. Les mots non vides vont être représentés par une matrice dont les colonnes représentent les mots et les lignes représentent les fréquences de ces mots pour dégager les mots les plus significatifs. Ainsi, les termes simples sont identifiés par l'élimination des mots vides de l'ensemble des mots qui composent le vocabulaire du corpus, [Baccari *et al.*, 12].

Par exemple pour la ligne suivante du document XML de la visioconférence « Chirurgie laparoscopique à Strasbourg » :

<ligne>The surgical act... consists of ... incision of few ... of millimeters</ligne>

Nous obtenons les :

- **Lmv** : The , of , few , ...
- **Lts** : surgical, act, consists, incision, millimeters, ...

IV.2.2.3. Pondération des termes simples

Dans cette étape, nous affectons à chaque terme un poids qui représente son pouvoir discriminant et son pouvoir représentatif dans le document où il apparaît. En effet, un terme ne représente d'une manière adéquate le document que si son degré d'importance dans ce document est significatif. Dans la littérature, nous distinguons deux types de pondération : locale et globale.

- La pondération locale

La pondération locale consiste à mesurer le pouvoir représentatif d'un terme dans un document du corpus, [Baziz, 05]. Elle utilise des informations locales du terme dans un document donné. Cette pondération est calculée de la manière suivante :

$$tf_{i,j} = \frac{n_{i,j}}{\sum_{k \in T_c} n_{k,j}} \quad (2.1)$$

où :

- $n_{i,j}$ est le nombre d'apparition du terme i dans le document j ;
- $n_{k,j}$ est le nombre d'apparition du terme k dans le document j . le dénominateur est le nombre d'occurrence de termes dans le document considéré ;
- T_c est l'ensemble des termes dans le corpus.

- La pondération globale

Cette pondération permet d'affecter à un terme une mesure reflétant son importance dans le corpus des documents. Un terme qui apparaît dans la majorité des documents est moins utile pour distinguer les documents les uns des autres, [Ben Aouicha, 09].

Dans notre approche, nous nous sommes intéressés à utiliser la pondération locale vue que l'indexation est faite document par document.

Les termes pondérés seront automatiquement ajoutés à la liste des termes simples. Pour les termes non pondérés, nous vérifions s'il s'agit des termes spécifiques ou non. Un mot spécifique est un mot de poids faible mais riche du côté sémantique. Il peut même avoir une seule occurrence dans la visioconférence à indexer tel que le mot « mini-caméra » qui figure une seule fois dans la visioconférence « chirurgie laparoscopique à Strasbourg ». Ce terme est considéré comme un terme spécifique pour la chirurgie laparoscopique car on ne peut pas effectuer ce type d'opération chirurgicale sans introduire cette mini-caméra dans l'appareil digestif du patient. Dans ce cas, et malgré son poids faible, le terme de poids faible est jugé spécifique. Donc, il sera ajouté à la liste des termes simples.

En se référant à la même visioconférence « Chirurgie laparoscopique à Strasbourg », nous trouvons que les termes « malade » et « patient » sont deux termes synonymes. Ils sont apparus 22 fois dans la visioconférence. Nous trouvons aussi que le nombre total des termes simples est 189 termes. D'où la pondération du terme est $22/189 = 0,116$. Ce poids est supérieur au seuil de fréquence fixé. Donc ces deux termes seront ajoutés à la liste des termes simples, [Zwidi *et al.*, 13].

IV.2.2.3. Algorithme d'extraction des termes simples

Algorithme Extraction des Termes Simples
<p>Entrée :</p> <p><i>DV</i> : Document xml de la visioconférence <i>DMV</i> : Document xml du mot vide <i>LP1</i> : {qui, lequel, laquelle, lesquels, ...} <i>LP2</i> : {celui-ci, ce dernier, celle-là, ceux-ci, ces derniers, ...} <i>LP3</i> : {il, elle, ils, elles}</p>
<p>Sortie :</p> <p><i>Lm</i> : Liste des mots du document <i>Lts</i> : Liste des termes simples</p>
<p>Variables :</p> <p><i>Lmv</i> : Liste des mots vides du document <i>m</i> : Un mot <i>doc</i> : Document <i>m1</i> : Un mot <i>nbm</i> : Nombre de mots non vides dans la visioconférence <i>occm</i> : Occurrence du mot <i>m</i> <i>ph</i> : Une phrase <i>phi</i> : i^{ème} phrase de la visioconférence <i>nph</i> : Nombre de phrases dans la visioconférence</p>
Début

```

// Prétraitement : décomposition du document en mots et phrases
// Suppression des mots vides de la visioconférence
Pour chaque phrase ph du doc faire
  Pour chaque mot m dans ph faire
    Si m appartient à Lmv alors
      Supprimer m du ph
    Fin si
  Fin pour
Fin pour
// Extraction des termes simples
Pour chaque phrase phi du doc faire
  Pour chaque mot m dans ph faire
    Ajouter m à Lm
     $occm \leftarrow 0$ 
     $nbm \leftarrow nbm + 1$ 
     $j \leftarrow i$ 
    Tant que ( $j \leq nph$ ) faire
      Pour chaque mot m1 de la phrase phj faire
        Si m et le premier terme de phj alors
           $p \leftarrow j + 1$ 
          Tant que ( $p \leq nph$  et le premier mot php appartient à LP3) faire
             $occm \leftarrow occm + 1$ 
            Supprimer le premier mot de php
             $p \leftarrow p + 1$ 
          Fin tant que
        Si non
          Si m est le premier mot de la phrase phj alors
            Si le premier mot de la phrase phj+1 appartient à LP2 alors
               $occm \leftarrow occm + 1$ 
              Supprimer le premier mot de la phrase phj+1
            Si non
              Si m1 suit directement m et m1 appartient à LP1 alors
                 $occm \leftarrow occm + 1$ 
                Supprimer m1 de la phrase phj
              Fin si
            Fin si
          Fin si
        Fin si
      Si  $m = m1$  alors
         $occm \leftarrow occm + 1$ 
        Supprimer m1 de la phrase phj
    Fin si
  Fin pour

```

Finsi
Fin pour
poids (m) ← occm / nbm
Si *poids (m) > seuil* **alors**
Ajouter *m* à *Lts*
Si non
Si *m* est un mot spécifique **alors**
Ajouter *m* à *Lts*
Fin si
Fin si
Fin Tant que
Fin pour
Fin pour
Fin

Algorithme 1 : Extraction des Termes Simples

IV.2.3. Extraction des Termes composés

IV.2.3.1. Extraction des termes composés basés sur l'information mutuelle

Parfois nous trouvons des mots non vides ayant une haute fréquence mais faible du côté sémantique. Par exemple, dans une visioconférence médicale, nous pouvons remarquer la répétition du mot simple « salle » et qui peut être pauvre en sémantique. Ce cas peut être considéré comme une limite d'indexation basée sur la fréquence des mots simples. Si ce mot est accompagné par le mot « opération » et lié entre eux par la préposition « de », le terme « d'opération », dans ce cas, joue le rôle en langue française d'un complément du nom. Ce nouveau mot composé « salle d'opération » porte une information riche sémantiquement.

Pour achever cette indexation basée sur les mots composés, nous allons prendre en compte les compléments des noms qui sont généralement liés par des propositions ou déterminants tels que « avec », « par », « à », « aux », « de », « des », etc. Ces derniers sont considérés comme des mots vides et sont déjà supprimés à l'étape d'extraction des termes simples. Les termes composés ont, par la suite, la forme Nom+Nom ou bien Nom+Adjectif avec un Nom qui peut être à son tour composé, [Baccari *et al.*, 12] et [Zwidi *et al.*, 13].

Donc, pour désigner un nouveau concept dans un domaine, le principe est d'éviter de créer un terme nouveau et ce qui engendrerait une explosion rapide du lexique, [Harrathi *et al.*, 05]. Ce nouveau terme, terme composé, est créé à partir de données lexicales préexistantes. Ces termes composés sont des combinaisons de deux ou plusieurs mots, [Smadja, 93]. Avec un nouveau concept, il n'y a pas de nouveaux termes mais il y a des nouvelles combinaisons des mots pour le désigner. Ces combinaisons sont des séquences de mots qui seront considérés comme des nouveaux termes.

Dans notre approche, nous adaptons la démarche de F. Harrathi, [Harrathi, 09]. Pour extraire les termes composés, nous utilisons un processus itératif et incrémental. Il permet de

découvrir de nouveaux termes à partir de ceux existants. Le processus procède à l'extraction de nouveaux termes à partir d'une liste initiale de termes connus en utilisant une mesure statistique : l'Information Mutuelle Adaptée (IMA). Nous partons de la liste des termes simples. Nous calculons par la suite la valeur de l'IMA de chaque couple de mots. Nous proposons de ne pas prendre en compte la fréquence d'un mot vide durant le calcul de l'IMA. Par exemple pour le terme « Salle de », la fréquence du mot vide « de » ne sera pas prise en compte et elle sera substituée par la valeur de la fréquence du terme simple « salle ». Durant le processus d'extraction des termes composés, le terme « salle de » est marqué comme étant un « terme de construction ». Ce terme est supprimé à l'itération suivante. Les couples des termes dont la valeur de l'IMA est inférieure à une valeur seuil sont acceptés comme des termes composés. Le processus s'arrête si à une itération aucun nouveau terme n'est extrait. Pour un couple de mots (m_i, m_j) , l'information mutuelle adaptée est calculée de la manière suivante :

$$IMA(t_i, t_j) = \left(\frac{2 * f(t_i, t_j)}{f(t_i) + f(t_j)} \right) \quad (2.2)$$

Tenons l'exemple du terme complexe « dissection aortique », le mot « dissection » apparait autant de fois que « aortique » et que « dissection aortique »

$$IMA(t_i, t_j) = \left(\frac{2 * f(t_i, t_j)}{f(t_i) + f(t_j)} \right)$$

$$\begin{aligned} IMA(dissection, aortique) &= \left(\frac{2 * (dis\ section, aortique)}{f(dis\ section) + f(aortique)} \right) \\ &= \left(\frac{2 * (dis\ section, aortique)}{2 * (dis\ section, aortique)} \right) \\ &= 1 \end{aligned}$$

Cette valeur « 1 » est la meilleure valeur de l'IMA dans le cas où la fréquence de chaque mot formant le terme composé est égale à la fréquence de ce terme composé dans le document.

Nous bénéficions du fait que les termes simples de chaque visioconférence sont stockés selon leurs ordres d'apparence. Pour cela, nous n'allons pas comparer chaque terme simple de la visioconférence à tous les autres termes simples qui la composent pour voir s'ils peuvent former ensemble des termes composés car cette méthode peut aboutir à des résultats incorrects. En plus, elle est trop coûteuse en terme du temps puisqu'il est inutile de vérifier la possibilité qu'un couple de terme forme un terme composé si ces derniers ne sont pas successifs c'est-à-dire l'un est loin de l'autre dans leur ordre d'apparence dans la visioconférence. Pour dépasser ces limites, nous calculons l'IMA de chaque couple de termes simples successifs (ou parfois un couple composé d'un terme composé et d'un terme simple). Ensuite, nous comparons la valeur de l'IMA obtenue pour chaque couple à un seuil. En cas où

cette valeur est supérieure au seuil, le couple est considéré comme un terme composé. Nous concaténons le couple en un seul terme. Puis, nous répétons le même processus avec ce nouveau terme et le terme simple qui le suit. Le processus s'arrête si le couple dernièrement traité ne forme pas un terme composé. Dans ce cas, le dernier terme composé trouvé est ajouté à la liste des termes composés pour la première fois en évitant ainsi le fait d'ajouter les termes de constructions puis leurs suppressions. Une fois nous n'avons pas parcouru tous les mots de la visioconférence, nous continuons à vérifier la possibilité d'avoir d'autres termes composés, [Zwidi *et al.*, 13].

Le processus d'extraction des termes composés que nous utilisons, donc, est composé de trois étapes :

1. **Initialisation** : dans cette étape, nous initialisons la liste des termes composés par le contenu de la liste des termes simples ;
2. **Découverte de nouveaux termes** : dans cette étape, nous calculons l'information mutuelle adaptée entre un élément de la liste des termes composés et un mot corpus ;
3. **Ajout de nouveaux termes** : dans cette étape, si nous trouvons une valeur de l'information mutuelle adaptée supérieure à un seuil donné, nous ajoutons les couples de mots à la liste des termes composés.

Par exemple pour la ligne suivante du document XML de la visioconférence « Chirurgie laparoscopique à Strasbourg » :

<ligne>The surgical act... consists of ... incision of few ... of millimeters</ligne>

Et en calculant la valeur d'IMA, nous obtenons :

- **Lts** : surgical, act, consists, incision, millimeters, ...
- **Ltc** : surgical_act

IV.2.3.2. Pondération des termes composés

À cette étape, nous calculons la pondération pour refléter l'importance du terme dans le document. Cette pondération dépend de trois facteurs : la fréquence du terme composé dans ce document, les pondérations des termes simples qui le composent et la longueur du terme composé. Pour la mesure proposée, nous prenons en compte ces trois facteurs. Les deux premiers facteurs sont représentés par la mesure classique $tf_{i,j}$. La pondération d'un terme composé est proportionnelle à sa longueur. Nous augmentons cette pondération par $1 - \frac{1}{longueur(i)}$.

La mesure $Ctf_{i,j}$ est donc exprimée en fonction de ces facteurs de la manière suivante :

$$Ctf_{i,j} = \left(1 - \frac{1}{longueur(i)}\right) * tf_{i,j} + \frac{1}{longueur(i)} * \sum_{k \in i} tf_{k,j} \quad (2.3)$$

où :

- i : un terme composé ;

- j : un document ;
- k : un terme simple ;
- $tf_{i,j}$: le taux d'occurrence du terme i dans le document j ;
- $longueur(i)$: le nombre de termes simples qui participent dans la construction du terme composé i ;
- $tf_{k,j}$: la pondération du terme k dans le document j ;

En se référant à la même visioconférence « Chirurgie laparoscopique à Strasbourg », nous trouvons que le terme « surgical_act » est apparu 3 fois dans la vidéo. Et nous trouvons que le nombre total des termes composés est 35 termes, d'où la pondération du terme « surgical_act » est $3/35 = 0,085$. Nous trouvons aussi que la pondération du terme « Surgical » est 0,015 (3/189) et la pondération du terme « act » est 0,026 (5/198).

D'où :

$$Ctf_{(\text{surgical_act}, \text{Vidéo chirurgie à Starsbourg})} = (1 - 1/2) * 0.085 + 1/2 * (0.026 + 0.015) = 0,063.$$

La pondération du terme composé « surgical_act » est donc 0,063.

Pour chaque terme composé non pondéré, nous vérifions s'il est un terme composé spécifique. Si c'est le cas, il est jugé spécifique et il sera ajouté à la liste des termes composés.

IV.2.3.3. Algorithme d'extraction des termes composés

Algorithme Extraction des Termes Composés
<p>Entrée :</p> <p>Lts : Liste des termes simples du document</p> <p>$Seuil_f$: Seuil de fréquence</p> <p>$Seuil_IMA$: Valeur du seuil de l'information mutuelle adaptée</p>
<p>Sortie :</p> <p>Ltc : Liste des termes composés du document</p>
<p>Variables :</p> <p>m : Un mot</p> <p>t : Un terme</p> <p>nbm : Nombre de termes dans la liste des termes simples</p> <p>$nmti$: Nombre de mots dans le terme ti</p> <p>tc : Terme composé</p> <p>$nmtc$: Nombre des mots dans un terme composé</p> <p>CTF : Fréquence d'un terme composé</p> <p>TF : Fréquence d'un terme simple</p>
<p>Début</p> <p>// Découverte des nouveaux termes</p> <p>$i \leftarrow 1$</p> <p>Tant que $i \leq nbm$ faire</p> <p style="padding-left: 20px;">$j \leftarrow i+1$</p> <p style="padding-left: 20px;">// Ajout de nouveaux termes</p>

```

Tc ← ti
Tant que (la valeur de l'IMA (ti, mj) > seuil_IMA) faire
  Concaténer (ti, «_», mj)
  j ← j+1
Fin tant que
Si (nmti ≥ 2) alors
  
$$CTF_{ij} = (1 - \frac{1}{long(t_i)} * TF_{ij} + \frac{1}{long(t_i)} \sum_{k \in i} TF_{kj}$$

  Si CTFij > seuil alors
    Ajouter tc à ltc
  Fin si
  Si tc est un mot spécifique alors
    Ajouter tc à la liste au ltc.
  Fin si
Fin si
  i ← j
Fin tant que
Fin

```

Algorithme 2: Extraction des Termes composés

Conclusion

Nous avons détaillé dans ce chapitre notre proposition pour la modélisation du contenu visioconférence. Le choix de spécifier deux formes de représentation (structurelle et sémantique) a pour objectif de mettre en œuvre l'ensemble des caractéristiques audiovisuelles du document visioconférence. Dans la représentation structurelle, nous modélisons les informations d'une manière générale telles que les informations temporelles, les descriptions événementielles, ... La structure d'une visioconférence peut se définir comme l'organisation des différents éléments de son contenu lui donnant sa cohérence et sa forme. Cette structure reflète l'intention première de son auteur et constitue donc une étape importante vers sa compréhension.

Par contre, dans la représentation sémantique, nous modélisons les informations sémantiques par type de média. Nous essayons pour chaque type de média de prendre en considération la richesse et l'expressivité sémantique d'un document visioconférence qui leurs accompagnent et qui peuvent être dégagées après une interprétation du contenu. L'interprétation des différentes descriptions, nous, permet de mettre en œuvre un modèle unique et représentatif du contenu visioconférence.

L'indexation d'un document consiste en un ensemble de concepts issus ou non d'un vocabulaire contrôlé. Dans notre contexte, ces concepts sont reliés par des relations sémantiques appartenant à l'ontologie du domaine. Toutefois la pertinence des informations que l'ontologie recèle dépend directement de leurs mises à jour. Ces mises à jour, qui consistent principalement en l'ajout de concepts et de relations, sont le plus souvent réalisées de façon manuelle, ce qui rend la tâche fastidieuse et le résultat subjectif.

Dans la partie suivante, nous décrivons notre modèle d'indexation des documents audiovisuels tout en détaillant les différents modules de ce modèle.

Chapitre 2 : Indexation du contenu des visioconférences médicales

L'objectif de l'indexation n'est pas pour identifier les informations contenues dans les documents mais plutôt de caractériser les connaissances associées à ces documents. L'indexation consiste par conséquent à extraire des informations synthétisant leurs contenus permettant la récupération de tout ou partie du document original.

Après la phase de modélisation, une phase d'indexation est obligatoire pour mieux exploiter et réutiliser le contenu d'un document.

Introduction.....	113
I. Techniques utilisées.....	114
I.1. Ontologie dense et thésaurus	114
I.2. Choix d'ontologie de domaine	114
I.2.1. OntoMénélas.....	115
I.2.2. UMLS	115
I.3. Formalisme de représentation.....	116
II. Module d'indexation	118
II.1. Extraction des Concepts	118
II.1.1. Choix de la ressource externe.....	122
II.1.2. Ambiguïtés des termes	123
II.1.3. Pondération des concepts	123
II.1.4. Algorithme d'extraction des concepts.....	124
II.2. Extraction des relations sémantiques entre les concepts	125
II.3. Mise à jour de la base d'index.....	126
III. Représentation sous forme d'un graphe unique	130
V. Réseaux des graphes	133
Conclusion	133

Introduction

Le SRI visioconférence doit être capable de retrouver des séquences dont la description ne coïnciderait pas exactement avec les termes employés pour formuler la requête. Imaginons qu'un utilisateur souhaite « retrouver les segments visioconférence dont le docteur Mourad Kassab parle d'une chirurgie cardiaque ». Une recherche effectuée sur cette description en utilisant les mots clés « Mourad Kasseb », « Chirurgie » et « Cardiaque » va bien retourner des segments ou au moins l'un de ces trois mots clés existe mais rien ne prouve que dans le ou les segment(s) retrouvé(s) existe bien un segment pertinent par rapport à la requête. Pour remédier à ce problème, nous avons proposé tout d'abord de mettre en place un modèle qui soit suffisamment riche pour pouvoir décrire le contenu d'un document visioconférence. Ce modèle se base sur une description du contenu visioconférence par des concepts et des relations conceptuelles en utilisant le formalisme des graphes conceptuels. Nous avons également proposé d'enrichir ce modèle par des connaissances externes dans une base afin d'améliorer le module d'indexation et de recherche. Ces données externes permettent d'enrichir les descriptions des requêtes utilisateurs et les descriptions utilisées pour indexer les documents. Elle doit permettre :

- de définir un modèle permettant de représenter et d'enrichir la sémantique des index de manière à ce que le système soit capable de faire des raisonnements. Par exemple, si une visioconférence est indexée par des concepts « chirurgien », « salle d'opération » et « cœur » en utilisant les ontologies, le système doit être capable d'inférer qu'il s'agit d'une chirurgie.
- de naviguer dans la structure des documents retrouvés et d'inférer le contexte.

Pour prendre en compte les besoins des précisions, nous proposons d'utiliser plusieurs ontologies de domaine pour faciliter l'échange des descriptions et de contribuer à améliorer la précision et l'efficacité du SRI visioconférence.

Ce chapitre est organisé de la façon suivante. La première section présente les différentes techniques utilisées dans notre proposition. Ces techniques comportent les ontologies utilisées et le modèle de graphe conceptuel. Ainsi, dans cette section, nous présentons les différentes ontologies utilisées dans notre système tout en justifiant nos choix.

Dans la deuxième section, nous décrivons les différentes étapes d'indexation. Nous nous basons sur le module d'analyse sémantique détaillé dans le chapitre précédent. Et à partir de la liste des termes (simples, composés et spécifiques) obtenus, nous extrayons les concepts tout en utilisant les ontologies choisies. Ensuite, nous cherchons les relations sémantiques qui peuvent exister entre les concepts.

Dans la troisième section, nous représentons les concepts et les relations sémantiques sous forme de graphes conceptuels. Enfin, les différents graphes obtenus durant la phase de modélisation structurelle et le module d'indexation seront fusionnés dans un seul graphe.

I. Techniques utilisées

I.1. Ontologie dense et thésaurus

Dans le cadre de nos travaux, nous nous intéressons plus particulièrement aux ontologies denses. Les ontologies denses sont des ontologies qui incluent les axiomes (les règles et contraintes) qui réagissent le domaine d'utilisation et intègrent les concepts et relations (structurés au sein de hiérarchies fondées sur la relation Spécialisation/Généralisation) du domaine considéré. Les axiomes ont pour but de définir la description des concepts et des relations permettant de représenter leur sémantique. Ils représentent les intentions des concepts et des relations du domaine et, de manière générale, les connaissances n'ayant pas un caractère strictement terminologique. Leur inclusion dans une ontologie peut avoir plusieurs objectifs : (1) définir la signification des composants, (2) définir des restrictions sur la valeur des attributs, (3) définir les arguments d'une relation, (4) vérifier la validité des informations spécifiées ou en déduire de nouvelles informations.

Pour assurer la cohérence entre les concepts ajoutés à notre système, il est indispensable qu'une certaine organisation et normalisation du vocabulaire choisi.

De même, l'utilisation d'un modèle de thésaurus permet de catégoriser des connaissances représentées dans la partie sémantique. Par exemple, le terme chirurgical peut représenter la catégorie professionnelle de deux chirurgiens orthopédistes « Mourad Kassab » et « Jean P.H. Duverseau ». Le thésaurus, en plus de l'utilisation des relations de spécialisation entre les concepts que nous pouvons trouver dans une taxonomie, utilise des relations linguistiques pour renforcer les relations entre les termes et élargir la sémantique du vocabulaire. Ces relations linguistiques peuvent être de relations de la synonymie, de la méronymie (relation de composition), l'hyponymie/hyperonymie (relation de spécialisation/généralisation) et de l'antonymie.

L'utilisation conjointe d'une ontologie dense et d'un thésaurus s'avère prometteuse dans le cadre d'un SRI documentaire par mots-clés. Il s'agit de décrire des ressources documentaires par un ensemble de métadonnées : {Sujet, Verbe, Objet} ou {Objet1, Relation, Objet2} selon le type de descripteur nécessaire, [Claude, 12]. Pour faciliter le module d'indexation des documents visioconférences et rendre des résultats plus cohérents, nous utiliserons le principe des ontologies denses et des thésauri où les concepts et les relations existantes dans ces ontologies sont présentées par le triplet {Objet, Relation, Objet}.

I.2. Choix d'ontologie de domaine

Considérant notre corpus de visioconférences, nous avons examiné certaines des terminologies médicales existantes. La recherche d'un terme donné peut donner plusieurs résultats selon l'ontologie utilisée. Par exemple, une recherche pour le concept *Carcinoma* doit considérer également les éléments indexés avec les termes *Malignant epithelial tumor* et *Epithelial Neoplasm* car l'ontologie SNOMED-CT nous précise que le premier terme est synonyme du *Carcinoma* et l'ontologie NCI Thesaurus nous précise que ce concept est un sous type (is_a) de l'*Epithelial Neoplasm*, [Jonquet et al., 09].

Pour résoudre ce problème, nous utiliserons, dans notre travail, deux ressources médicales externes telles que : OntoMénélas et UMLS, sachant que UMLS est la fusion de plusieurs ressources telles que MeSH, SNOMED-CT et RXNORM. Dans ce qui suit, nous donnerons un aperçu sur ces ressources.

I.2.1. OntoMénélas

OntoMénélas nous permet dans un premier temps d'identifier les concepts associés aux termes et dans un deuxième temps d'identifier les relations entre ces concepts.

L'ontologie OntoMénélas a été développée durant le projet MÉNÉLAS par Bruno Bachimont, Pierre Dujols, Jacques Bouaud, Jean Charlet et Pierre Zweigenbaum. Cette ontologie couvre le domaine médical plus précisément dans le domaine de chirurgie cardiaque. Cette ressource a été créée dans le but de faciliter la recherche et l'intégration d'informations.

OntoMénélas est formé de deux composantes principales :

1. **La top-ontologie** : il regroupe principalement les concepts (1500 concepts) et les termes associés à ces concepts. Ces termes sont écrits dans une ou plusieurs langues. Ces concepts et termes sont issus de différentes ressources sémantiques. Des variations syntaxiques et lexicales des termes sont parfois données.
2. **Le réseau sémantique** : il définit l'organisation des concepts et les relations entre ces concepts. Dans OntoMénélas, les concepts sont organisés en classe. À chaque concept au moins une classe est associée. Ces classes forment des types sémantiques. Ces types sont reliés entre eux par des relations sémantiques (500 relations).

OntoMénélas est l'une des meilleures ressources sémantiques pour indexer des documents audiovisuels couvrant le domaine de chirurgie cardiaque, [Maisonasse, 08].

Dans la ressource sémantique, OntoMénélas est identifié par un identificateur unique (rdf:ID =). Chaque concept possède un ou plusieurs labels qui sont des termes associés à ce concept. Un concept est relié à une ou plusieurs chaînes de caractères (hiddenLabel) qui dénotent les concepts. Les hiddenLabel seront utilisés dans notre méthode pour identifier les concepts associés aux termes (rdf:ID). Ils sont liés à une langue (xml:lang=) . L'extrait suivant montre un exemple de concept dans OntoMénélas (cf. **Extrait 1**).

I.2.2. UMLS

UMLS est un méta-thésaurus qui résulte de la fusion de différentes sources (thésaurus, liste d'autorités). Même s'il n'est ni complet ni constant, il contient plus de 1 millions de concepts reliés à plus de 5,5 millions de termes dans 17 langues. UMLS ne constitue cependant pas une ontologie au sens strict du terme, car aucune description formelle des concepts n'est fournie. UMLS définit plutôt des groupes de termes, chaque groupe, identifié à un concept, étant constitué d'un ou plusieurs termes et de leurs variantes, [Maisonasse *et al.*, 09].

L'UMLS est constitué de trois composantes principales connaissances, [Drame, 14] et [Le Maguer *et al.*, 15] :

```

<owl:Class rdf:ID="HeartRhythm">
  <DOE:prefLabel xml:lang="fr">rythme cardiaque</DOE:prefLabel>
  <DOE:prefLabel xml:lang="en">heart rhythm</DOE:prefLabel>
  <rdfs:subClassOf>
    <owl:Class rdf:about="#PhysiologicalAttribute"/>
  </rdfs:subClassOf>
  <DOE:hiddenLabel xml:lang="fr">RythmeCardiaque</DOE:hiddenLabel>
  <DOE:hiddenLabel xml:lang="en">HeartRhythm</DOE:hiddenLabel>
  <DOE:altLabel xml:lang="en">HeartRhythm</DOE:altLabel>
  <DOE:definition xml:lang="en">OG Representation :
heart_rhythm: _x]-
  (attr)&lt;--[heart]
  (val_qual)--&gt;[heart_rhythm_val]
  #</DOE:definition>
</owl:Class>

```

Extrait 1: Exemple de concept dans OntoMénélas

- Le Meta-thesaurus des concepts médicaux (prises de vocabulaires établis tels que CIM-9-CM et MeSH) et inter-concept relations dont l'objectif principal est de cartographier entre les systèmes de codage et donc assurer un échange d'informations entre les différentes bases de données cliniques et des systèmes.
- Le réseau sémantique, qui spécifie les catégories à laquelle les concepts médicaux définis dans le Meta-thesaurus, peut appartenir, et les relations sémantiques qui peuvent être attribuées entre ces concepts et leurs attributs. 135 types sémantiques et 54 relations sémantiques existent, [Tilley *et al.*, 04].
- Le lexique spécialisé qui contient des informations syntaxiques, morphologiques et orthographiques pour les mots médicaux et communes dans la langue anglaise. Le lexique et ses ressources lexicales associées sont utilisés pour générer les index des Meta-thesaurus et aussi avoir une large applicabilité dans les applications de traitement naturel dans le domaine médical, [Tilley *et al.*, 04].

I.3. Formalisme de représentation

L'indexation est donc basée sur des formalismes ou modèles de représentation de connaissances comme les réseaux sémantiques, les graphes conceptuels et les graphes sémantiques.

La modélisation des connaissances cherche à représenter un ensemble de données ou connaissances sous une forme adaptée pour qu'un opérateur humain et/ou une machine, puisse les interpréter et les manipuler. Une représentation est définie selon un modèle qui

fournit les règles syntaxiques de modélisation, appelées la syntaxe. Le modèle peut être muni d'une sémantique, logique par exemple, pour définir clairement le sens de ce qui est modélisé. Le modèle est dit formel si les modélisations basées sur ce modèle sont sémantiquement sans ambiguïté, [Djamel *et al.*, 08].

La formalisation de connaissance commence par une conceptualisation qui consiste en un ensemble d'objets, concepts et autres entités sur lesquelles les connaissances sont exprimées (appelé univers de discours) et les relations entre eux.

Cette idée permettant d'établir des relations entre des objets et de les représenter dans la base de connaissances d'un système informatique est très fréquente. Cependant, la plupart des formalismes de représentation de connaissances n'exploitent pas toute la richesse de la sémantique de ces relations, [Gandon, 08].

La recherche de documents structurés ou complexes exige plus que jamais, outre un langage d'indexation robuste et expressif, la prise en charge de la sémantique des relations ainsi que leurs propriétés. À travers une étude des nouvelles exigences auxquelles la recherche d'informations d'aujourd'hui doit répondre, nous proposons un modèle d'indexation relationnel pour les documents. Les connecteurs sémantiques sont vus comme des relations permettant de construire des expressions nouvelles représentant des concepts nouveaux ou des situations nouvelles.

Une caractéristique intéressante de ce modèle est qu'il conduit à étendre certains formalismes de représentation de connaissances par des notions utiles en recherche d'information

La représentation des connaissances selon Wikipédia français désigne un ensemble d'outils et de technologies destinés d'une part à représenter et d'autre part à organiser le savoir humain pour l'utiliser et le partager.

La modélisation des connaissances est un sujet majeur pour différentes problématiques. Elle, en tant que processus intermédiaire entre l'analyse et l'exploitation, fait de plus en plus appel à des langages de représentation.

Différentes méthodes, formalismes et outils ont été proposés et mis en œuvre pour modéliser, stocker et utiliser les informations au sein des bases de données ou des bases de connaissances pour répondre aux besoins d'un utilisateur jouant par la suite un rôle central dans l'interaction et le dialogue avec l'utilisateur tels que les graphes conceptuels et les ontologies.

La conception d'un modèle d'indexation des visioconférences basé sur les graphes conceptuels et les différents points reliés au modèle nécessite un ensemble des besoins à satisfaire.

Ces besoins peuvent être regroupés selon deux aspects. Le premier est dit aspect utilisateur qui spécifie, au moyen d'une requête, un besoin d'information. Ce besoin doit se traduire d'une façon flexible et claire. Ceci peut être réalisé en prenant en considération tous les types de média, des différents niveaux d'analyse (signal, sémantique) et également du contexte d'où la nécessité d'une bonne modélisation des requêtes pouvant être formulées en utilisant les différents types de média.

Le deuxième aspect est dit aspect système. Le système proposé doit traiter les requêtes puis présenter efficacement les résultats. Pour achever ceci, il est nécessaire de modéliser efficacement les documents. Dans notre approche, nous choisissons le formalisme des graphes conceptuels pour la modélisation des connaissances en raison de son adaptation aisée à l'approche logique de la recherche d'information notamment dans le cas des documents multimédia.

II. Module d'indexation

L'indexation est une étape qui consiste à analyser le document lors de l'organisation du fond documentaire afin de produire un ensemble de mots clés, appelés aussi « *descripteurs* », que le système pourra gérer aisément puis utiliser dans le module de recherche ultérieur.

Ce module permet d'extraire les concepts des documents de base. Dans un SRI, un document est considéré comme un support qui véhicule de l'information. Le résultat de ce module d'indexation est représenté sous forme de liste de descripteurs. Cette représentation est appelée un index de document.

Les ontologies améliorent considérablement la pertinence des résultats dans la recherche de documents audiovisuels. C'est la raison pour laquelle nous avons opté pour la méthode d'indexation à l'aide de l'ontologie, [Zwidi *et al.*, 13].

Cette amélioration est due au fait que le module d'indexation prend en compte les concepts ainsi que les différentes relations entre eux. Par conséquent et contrairement aux méthodes d'indexation par mots clés simples et statiques qui s'intéressent à l'existence ou la non existence d'un mot dans un document, cette méthode prend en compte la sémantique des termes recherchés, [Hamadi *et al.*, 15].

Le module d'indexation d'un document audiovisuel (DAV) est illustré à travers la **figure 2.13**.

Le critère principal, dans l'extraction des descripteurs, doit toujours être la valeur potentielle d'un concept en tant qu'élément dans l'expression du contenu d'un document et dans sa recherche d'information.

Le module d'indexation permet de générer automatiquement les concepts ascendants des concepts extraits. Ce qui permettra, par exemple, à partir du concept « specialist » de générer les concepts « doctor », « cardiologist » comme descripteurs du document bien qu'ils ne soient pas présents dans son contenu.

De même, il consiste à identifier les concepts dénoté par les termes précédemment extraire (dans la section III.2 du chapitre précédent). Le troisième est l'extraction des relations entre les concepts.

II.1. Extraction des Concepts

Durant cette étape, nous extrayons les concepts à partir des visioconférences médicales. Ces concepts sont dénotés dans les documents XML par des termes simples ou composés. Ces termes ont été extraits pendant les étapes précédentes. Pour effectuer la correspondance entre les termes et les concepts associés à ces termes, nous utilisons une ou

deux ressources externes de domaine médical (OntoMénélas, UMLS) sachant qu'une ressource est composée d'un ensemble de concepts C et un ensemble de relations entre les concepts R , [Derbas, 14]. En étudiant ces ressources, nous avons constaté qu'une ressource ne contient pas tous les concepts du domaine médical. Pour cette raison, nous avons choisi la combinaison d'une ontologie et un méta-thésaurus pour couvrir tout le domaine médical.

Ainsi, la structure d'une ontologie est définie par $O = \{C, R, \leq^c\}$

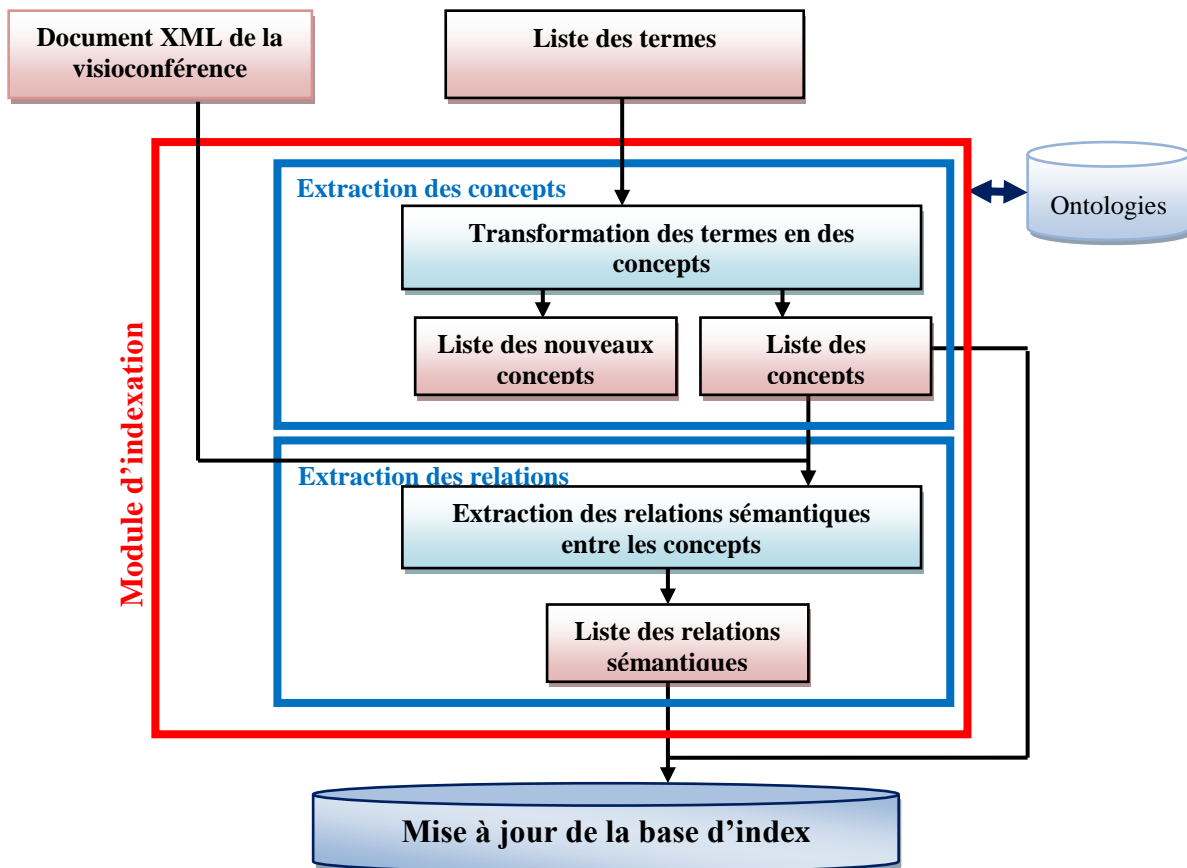


Figure 2. 13 : Module d'indexation

Où

- C : l'ensemble des concepts de l'ontologie ;
- R : l'ensemble des relations entre les concepts de l'ontologie ;
- \leq^c : $C \times C$ est un ordre partiel sur C , il définit la hiérarchie de concepts ; $\leq^c (c_1, c_2)$ signifie que c_1 subsume c_2 (relation orientée).

Dans une ontologie, chaque concept est identifié par un identifiant unique. À chaque concept, un ou plusieurs termes ont été associés. Ces termes sont appelés « labels » qui se divisent en label « préféré » et label « alternatif ». Les labels alternatifs sont considérés comme des

synonymes des labels préférés, [Chebil *et al.*, 15]. Par exemple, le concept « C1516400 » de l'UMLS possède un label préféré et deux labels alternatifs (cf. **Figure 2.14**).

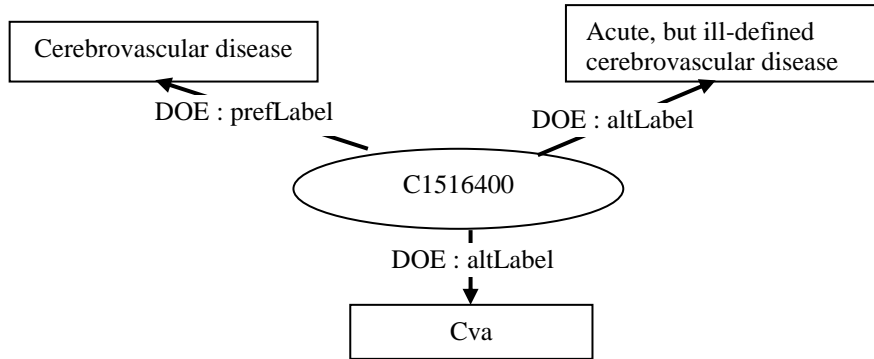


Figure 2. 14 : Exemple d'un concept décrit par DOE

Le concept de la figure ci-dessus est décrit de la manière suivante :

```

<rdf:RDF>
<owl:Class rdf:ID="C1516400">
<DOE:prefLabel>Cerebrovascular disease</DOE:prefLabel>
<DOE:altLabel>Acute, but ill-defined cerebrovascular disease</DOE:altLabel>
<DOE:altLabel> Cva</DOE:altLabel>
</owl:Class>
</rdf:RDF>
  
```

Extrait 2: Extrait XML du concept C1516400

Dans l'ontologie, un ensemble de termes est utilisé afin de labéliser les concepts et les relations entre les concepts. Cet ensemble forme le vocabulaire de l'ontologie et sera noté : V_o dont $V_o = \{V_{OC}, V_{OR}\}$

Où

- V_{OC} : l'ensemble des termes utilisés pour dénoter les concepts de l'ontologie ;
- V_{OR} : l'ensemble des termes utilisés pour dénoter les relations de l'ontologie.

En utilisant l'ensemble V_{OC} , nous définissons l'opérateur de référence de terme S_c . Cet opérateur permet de déterminer le(s) concept(s) dénoté(s) par un terme. Un opérateur de référence de concept L_c est défini. Ces deux opérateurs sont déterminés par les équations suivantes :

$$\left\{ \begin{array}{l} \forall t \in V_{oc}, S_c(t) = \{c \in C / c \text{ est dénoté par } t\} \quad \text{et} \\ \forall c \in C, L_c(c) = \{t \in V_{oc} / t \text{ dénote } c\} \end{array} \right. \quad (2.4)$$

En se référant au même exemple mentionné avant, nous obtenons pour le concept « C1516400 » :

$$\left\{ \begin{array}{l} \mathcal{S}_c ("Cerebrovascular\ disease") = \{ "C1516400" \} \text{ et} \\ \mathcal{S}_c ("Acute, but ill-defined cerebrovascular disease") = \{ "C1516400" \} \text{ et} \\ \mathcal{S}_c ("Cva") = \{ "C1516400" \} \text{ et} \\ L_c ("C1516400") = \{ "Cerebrovascular\ disease", "Acute, but ill-defined \\ cerebrovascular disease", "Cva" \} \end{array} \right.$$

Vu que dans nos travaux, nous utilisons deux ressources externes en même temps, nous associons à chaque concept un ensemble de labels préférés traduites en anglais et en français, (cf. **Figure 2.15**).

Nous ajoutons à la balise du label, un tag « Xml : Onto ». Ce tag permet d'identifier la ressource utilisée. En parcourant l'ontologie OntoMénélas, nous avons trouvé que le concept **C1516400** de l'UMLS correspond au concept **HeartRhythm** dans OntoMénélas. D'où l'extrait 2 sera comme présenté dans l'extrait 3.

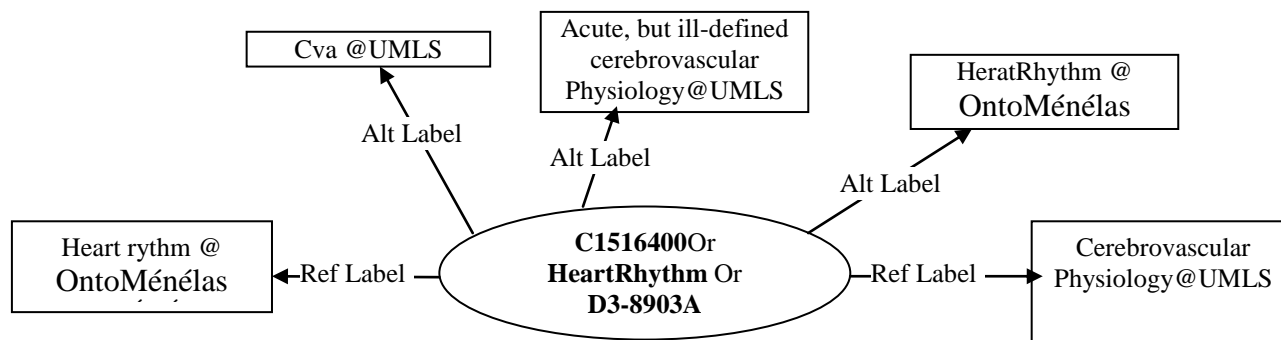


Figure 2. 15 : Exemple de concept d'une ontologie

```
<rdf:RDF>
<owl:Class rdf:ID="C1516400">
<DOE:prefLabel Xml:Onto = "UMLS"> Cerebrovascular Physiology</DOE:prefLabel>
<DOE:altLabel Xml:Onto = "UMLS">Acute, but ill-defined cerebrovascular Physiology</DOE:altLabel>
<DOE:altLabel Xml:Onto = "UMLS"> Cva </DOE:altLabel>
<DOE:prefLabel Xml:Onto = "OntoMénélas"> Heart rhythm </DOE:prefLabel>
<DOE:altLabel Xml:Onto = "OntoMénélas"> HeratRhythm </DOE:altLabel>
</DOE:altLabel>
</owl:Class>
</rdf:RDF>
```

Extrait 3: Extrait XML du concept C1516400

Après la projection des termes sur l'ontologie, nous trouvons que chaque terme peut être associé à un ensemble de concepts. À titre d'exemple, pour un terme T1 nous pouvons trouver les concepts C1, C3 et C4, pour un terme T2 nous pouvons extraire les concepts C2, C4, C3 et C5 et pour un terme T3 nous pouvons dégager les concepts C1, C2 et C6. Soit le tableau suivant présentant les concepts dégagés pour chaque terme :

Tableau 2.2 : Concepts associés

Termes	Concepts
T1	C1, C3, C4
T2	C2, C4, C3, C5
T3	C1, C2, C6

Or notre indexation est faite par concept et pas par terme. Nous représentons, ainsi, chaque concept trouvé en fonction des termes qui lui sont associés. À partir de l'exemple donné ci-dessus, nous obtenons le **tableau 2.3**.

La méthode que nous proposons pour l'extraction des concepts à partir d'un document XML d'une visioconférence médicale consiste à affecter à chaque terme d'un document les concepts associés. Afin d'identifier les concepts associés à chaque terme, nous utilisons les relations Sc définies précédemment.

Tableau 2.3 : Concepts dénotés

Concepts	Termes
C1	T1, T3
C2	T 2, T3
C3	T1, T2
C4	T1, T2
C5	T2
C6	T3

II.1.1. Choix de la ressource externe

Pour l'indexation sémantique basée sur les concepts, les descripteurs décrivant les concepts sont représentés par les termes trouvés dans le module de modélisation. Ces termes

seront projetés sur une (des) ressource(s) sémantique(s) externe(s) pour identifier les concepts qui y sont associés ainsi que les relations entre eux. Les ressources externes (OntoMénélas et UMLS) utilisées sont jugées parmi les ressources sémantiques les plus utiles pour l'indexation des documents audiovisuels couvrant le domaine médical.

II.1.2. Ambigüités des termes

Le problème d'ambigüité des termes se pose lors de l'association des termes aux concepts. Deux types d'ambigüité existent : une ambigüité langagière et une ambigüité sémantique, [Adbulahhad *et al.*, 11a], [Adbulahhad *et al.*, 11b], [Drame, 14] et [Chifu, 15].

II.1.2.1. Ambigüité langagière

Ce problème est rencontré dans le cas des documents multilingues. Il s'agit de trouver deux termes appartenant à des langues différentes mais ayant la même forme dans un texte. Ce type d'ambigüité ne sera pas traité dans nos travaux vu que nous utilisons des documents monolingues.

II.1.2.2. Ambigüité sémantique

L'ambigüité sémantique existe lorsque nous trouvons plusieurs concepts dénotés par le même terme (un terme peut être le label de plusieurs concepts dans l'ontologie). Pour résoudre ce problème, nous cherchons un autre concept C' en relation avec le concept C dénoté par le terme ambigu t dans l'ontologie. Si nous trouvons le concept C' , nous considérons le concept C comme étant le concept dénoté par le terme t . Si non, nous utilisons une autre ressource externe. Dans le cas où nous n'avons trouvé aucun concept en relation avec le concept C (en utilisant les deux ressources), nous prenons tous les concepts dénotés par le terme considéré.

II.1.3. Pondération des concepts

Durant cette étape, nous pondérons la liste des concepts dégagés. Un concept est représenté dans la visioconférence par un ou plusieurs termes dont la fréquence de chacun est déjà calculée lors de l'extraction des termes simples et des termes composés. La Fréquence d'un Concept (CF) est égale à la moyenne des fréquences des termes qui le présentent dans le document, [Zwidi *et al.*, 13].

$$FC_{ij} = \frac{1}{n} * \sum_{i=1}^n FT_{ij} \quad (2.5)$$

- FC_{ij} : fréquence d'un concept i dans le document j
- n : le nombre de termes jugés représentatifs du concept i dans le document j
- $\sum_{i=1}^n FT_{ij}$: somme des fréquences des termes jugés représentatifs du concept i dans le document j

La pondération des concepts cherche de trier proportionnellement les concepts à leurs importances dans le document visioconférence. Le concept jugé plus important que les autres est celui ayant la fréquence la plus élevée. Nous allons, alors, les ordonner selon leurs fréquences de plus fréquents au moins fréquents.

II.1.4. Algorithme d'extraction des concepts

Le processus d'extraction des concepts est composé de deux étapes. Dans la première, nous ne traitons que les termes non ambigus. Ces termes seront utilisés par la suite pour désambiguïser les termes ambigus dans la deuxième étape.

L'algorithme de la méthode d'extraction des concepts est le suivant :

Algorithme Extraction des concepts
<p>Entrée</p> <p><i>Lts</i> : Liste des termes simples</p> <p><i>Ltc</i> : Liste des termes composés</p> <p><i>Lco</i> : Liste des concepts de l'ontologie</p>
<p>Sortie</p> <p><i>Lcp</i> : Liste des concepts pondérés</p>
<p>Variables</p> <p><i>Lt</i> : Liste des termes // <i>formée par les termes simples et les termes composés</i></p> <p><i>t</i> : Un terme</p> <p><i>c</i> : Un concept</p> <p><i>DV</i> : Document XML de la visioconférence</p>
<p>Début</p> <p>// <i>initialisation de la liste des termes</i></p> <p>$Lt \leftarrow Lts \cup Ltc$</p> <p>// <i>identification des concepts associés aux termes non ambigus</i></p> <p>Pour chaque terme <i>t</i> dans <i>DV</i> faire</p> <p> Identifier les concepts associés à <i>t</i> // <i>nous utilisons Sc</i></p> <p> Si <i>t</i> n'est pas ambigu alors</p> <p> Ajouter (<i>DV</i>, <i>C</i>, poids (<i>t</i>, <i>DV</i>)) à <i>Lcp</i> // <i>c est le concept identifié</i></p> <p> Finsi</p> <p>Fin pour</p> <p>// <i>identification des concepts associés aux termes ambigus de point de vu sémantique</i></p> <p>Pour chaque terme <i>t</i> ambigu dans <i>DV</i> faire</p> <p> Identifier les concepts associés à <i>t</i></p> <p> Chercher dans <i>DV</i> un <i>CI</i> dénoté par le terme <i>tI</i> qui apparait dans une même phrase que <i>t</i></p> <p> Si <i>CI</i> existe alors</p> <p> $C \leftarrow$ l'ensemble des concepts dénotés par <i>t</i> qui sont en relation dans l'ontologie avec le concept <i>CI</i></p> <p> Ajouter (<i>DV</i>, <i>C</i>, poids (<i>t</i>, <i>DV</i>)) à <i>Lcp</i> // <i>c est le concept identifié</i></p> <p> Sinon</p> <p> Ajouter (<i>DV</i>, <i>C</i>, poids (<i>t</i>, <i>DV</i>)) à <i>Lcp</i> // <i>C est l'ensemble des concepts associés à t</i></p> <p> Finsi</p>

Fin pour
Fin

Algorithme 3: Extraction des concepts

II.2. Extraction des relations sémantiques entre les concepts

Afin d'extraire les relations sémantiques entre les concepts extraits dans la section précédente, nous nous basons sur les ressources sémantiques utilisées, dans notre cas une ontologie et un méta-thésaurus. Ces relations sont définies dans les ressources par les types relations. Nous admettons l'hypothèse mentionnée dans, [Maisonasse, 08] : « une relation existe entre deux concepts d'un document si ces deux concepts apparaissent dans la même phrase et si la ressource sémantique définit cette relation sémantique ». Parmi les relations sémantiques, nous trouvons à titre d'exemple la relation d'identité, de synonymie, de classification sorte_de, d'homonymie, d'équivalence et d'autonymie.

Si nous prenons l'extrait 3 de XML d'une visioconférence médicale et en utilisant UMLS, nous détectons les concepts **C0334046** et **C1302773** dénotés par les termes « *mild dysplasia* » et « *low grade squamous intraepithelial lesion* ».

En appliquant l'hypothèse de Maisonasse, nous trouvons que les concepts **C0334046** et **C1302773** appartiennent à la même phrase, et que ces deux concepts sont reliés par la relation « *is_finding_of_disease* ». En utilisant UMLS, nous trouvons que cette relation est définie comme une relation sémantique. C'est la relation **R54390434** d'UMLS.

```
<? XML version = "1,0" encoding "iso-8859-1" standalone = "Yes"?;>
<DOC>
<ID> 0062782 </ID>
<Diagnostic> mild dysplasia of squamous epithelium CIN I, LSIL ..
(6278)
Low grads squamous intraepithelial lesion, coilocyte. </Diagnostic>
<Description> Atypical cells corresponding to a mild dysplasia. Small air bubbles.
</Description>
</Doc>
```

Extrait 4 : Extrait XML d'une visioconférence médicale

Notre algorithme d'extraction des relations sémantiques entre les concepts est le suivant :

Algorithme Extraction des relations sémantique entre concepts

Entrée

Lc : Liste des concepts

Lro : Liste des relations de l'ontologie

Lt : Liste des termes de la visioconférence

Sortie

<i>Lrs</i> : Liste des relations sémantiques
<p><u>Variables</u></p> <p><i>ph</i> : Une phrase</p> <p><i>Lcph</i> : Liste des concepts d'une phrase <i>ph</i></p> <p><i>ci, cj</i> : Concepts appartenant à <i>Lc</i></p> <p><i>r</i> : Relation</p> <p><i>DV</i> : Document XML de la visioconférence</p>
<p>Début</p> <p>Pour chaque phrase <i>ph</i> dans <i>DV</i> faire</p> <p>// <i>extraire la liste des concepts inclus dans la phrase Lcph</i></p> <p>Pour chaque couple de concepts <i>ci</i> et <i>cj</i> dans <i>ph</i> faire</p> <p>Pour chaque terme <i>t</i> de la liste des termes <i>Lt</i> de <i>ph</i> et <i>t</i> n'appartient pas à <i>Lcph</i> faire</p> <p>$r \leftarrow t$</p> <p>Si appartient à <i>Lro</i> alors</p> <p>ajouter <i>r</i> à <i>Lrs</i></p> <p>Fin si</p> <p>Fin pour</p> <p>Fin pour</p> <p>Fin pour</p> <p>Fin</p>

Algorithme 4: Extraction des relations sémantiques entre concepts

II.3. Mise à jour de la base d'index

Dès qu'un nouveau document arrive au système, la mise à jour de l'index est réalisée selon l'algorithme 5. L'index est mis à jour en reliant le document ajouté avec les concepts de l'ontologie et en sauvegardant la fréquence d'apparition de chaque concept. Cette fréquence sera utile pour le tri des documents trouvés lors du module de recherche.

Algorithme Mise à jour d'index : Ajout document
<p><u>Entrée</u></p> <p><i>DI</i> : Un nouveau document</p>
<p><u>Sortie</u></p> <p><i>Index</i> : Index mis à jour</p>
<p><u>Variables</u></p> <p><i>Ci</i> : Concept associé au niveau document</p> <p><i>Cj</i> : Concept associé à l'index</p> <p><i>Ri</i> : Relation sémantique</p> <p><i>Pi</i> : Poids du concept <i>Ci</i></p> <p><i>NbDoc</i> : Nombre de documents correspondants à <i>Ci</i></p>

Début

Changer l'état du document DI à ajouter

Pour chaque concept C_i du document DI **faire**

Si C_i n'existe pas dans $Index$ **alors**

Ajouter C_i

Relier C_i à C_j correspondant en utilisant la relation R_i

Sauvegarder pour C_i son poids P_i dans $Index$

Incrémenter $NbDoc$ correspondant à C_i

Si non

Incrémenter $NbDoc$ correspondant à C_i

Sauvegarder pour C_i son poids P_i dans $Index$

Fin si

Fin pour

Fin

Algorithme 5: Prise en compte de l'ajout d'un nouveau document

Pour un document supprimé, son statut est changé en « effacé » avant de mettre à jour les différentes informations relatives au document dans la base. Donc, ce document ne sera pas pris en compte par les requêtes. Les concepts correspondants à ce document doivent être supprimés s'ils ne déterminent pas d'autres documents. La suppression du document est faite par l'algorithme 6.

Algorithme Mise à jour d'index : suppression document

Entrée

DI : Document à supprimer

Sortie

$Index$: Index mis à jour

Variables

C_i : Concept associé au niveau document

C_j : Concept associé à l'index

R_i : Relation sémantique

P_i : Poids du concept C_i

$NbDoc$: Nombre de documents correspondants à C_i

Début

Changer l'état du document DI à l'état « Effacé »

//le document sera ignoré par toutes les requêtes

Pour chaque concept C_i du document DI **faire**

Décrémenter $NbDoc$ correspond à C_i

Si $NbDoc = 0$ **alors**

Supprimer C_i

Supprimer R_i reliant C_i à C_j

Fin si

Fin pour

Fin

Algorithme 6: Prise en compte de la suppression d'un nouveau document

Lors d'une évolution d'un document dans le temps, des ajouts, des modifications ou des suppressions d'information peuvent être faites. Ces événements nécessitent la mise à jour des index, pour garder la cohérence entre l'index et les documents du corpus, car ils peuvent provoquer l'apport de nouveaux concepts, de suppression des concepts ou de modification des fréquences d'apparition des concepts.

L'ajout de nouvelles informations dans un document peut engendrer les événements suivants :

- indexation par d'autres concepts des ressources externes ;
- indexation par un nouveau concept dans le document ;
- indexation par une nouvelle relation reliant des concepts.

Algorithme Mise à jour d'index : modification document – ajout information

Entrée

DI : Document modifié

MI : Partie modifiée du document

Sortie

$Index$: Index mis à jour

Variables

C_i : Concept associé à MI

C_j : Concept associé à DI

C_k : Concept associé à $Index$

R_i : Relation sémantique

P_i : Poids du concept C_j

Début

Pour chaque concept C_i du MI **faire**

Si C_i n'existe pas dans la liste de C_j **alors**

Si C_i n'existe pas dans $Index$ **alors**

Ajouter C_i

Relier C_i à C_j ou C_k par R_i

Mettre à jour P_i correspondant à C_i

Si non

Mettre à jour P_i correspondant à C_i

Fin si

Si non

Incrémenter P_i
Fin si
Fin pour
Fin

Algorithme 7: Modification d'un document : ajout d'information

La suppression d'information dans le document peut générer les événements suivants :

- suppression des relations entre les concepts existantes dans les informations supprimées ;
- suppression des concepts.

La mise à jour de l'index sera traitée par l'algorithme 8.

Algorithme Mise à jour d'index : modification document – suppression information
<p><u>Entrée</u></p> <p>DI : Document modifié MI : Partie supprimée du document</p>
<p><u>Sortie</u></p> <p>$Index$: Index mis à jour</p>
<p><u>Variables</u></p> <p>C_i : Concept associé à MI C_j : Concept associé à DI C_k : Concept associé à $Index$ R_i : Relation sémantique P_i : Poids du concept C_j $NbDoc$: Nombre de documents associés à C_i</p>
<p>Début</p> <p>Pour chaque concept C_i du MI faire</p> <p style="padding-left: 20px;">Décrémenter P_i</p> <p style="padding-left: 20px;">Si $P_i = 0$ alors</p> <p style="padding-left: 40px;">Décrémenter $NbDoc$</p> <p style="padding-left: 40px;">Si $NbDoc = 0$ alors</p> <p style="padding-left: 60px;">Supprimer C_i</p> <p style="padding-left: 60px;">Supprimer R_i reliant C_i à C_j ou C_k</p> <p style="padding-left: 20px;">Fin si</p> <p style="padding-left: 20px;">Fin si</p> <p>Fin pour</p> <p>Fin</p>

Algorithme 8: Modification d'un document : suppression d'information

La modification des informations dans une séquence audiovisuelle se traduit par :

- suppression d'information ;
- ajout d'information.

Ceci explique les opérations suivantes :

- suppression des concepts ;
- suppression des relations ;
- indexation par un autre concept de l'ontologie ;
- indexation par une relation reliant deux concepts.

Dans tous les cas, la mise à jour de l'index n'est réalisée que sur le document concerné. Cela permet de diminuer le temps d'indexation tout en conservant la cohérence entre le corpus et l'index pour trouver les documents les plus pertinents.

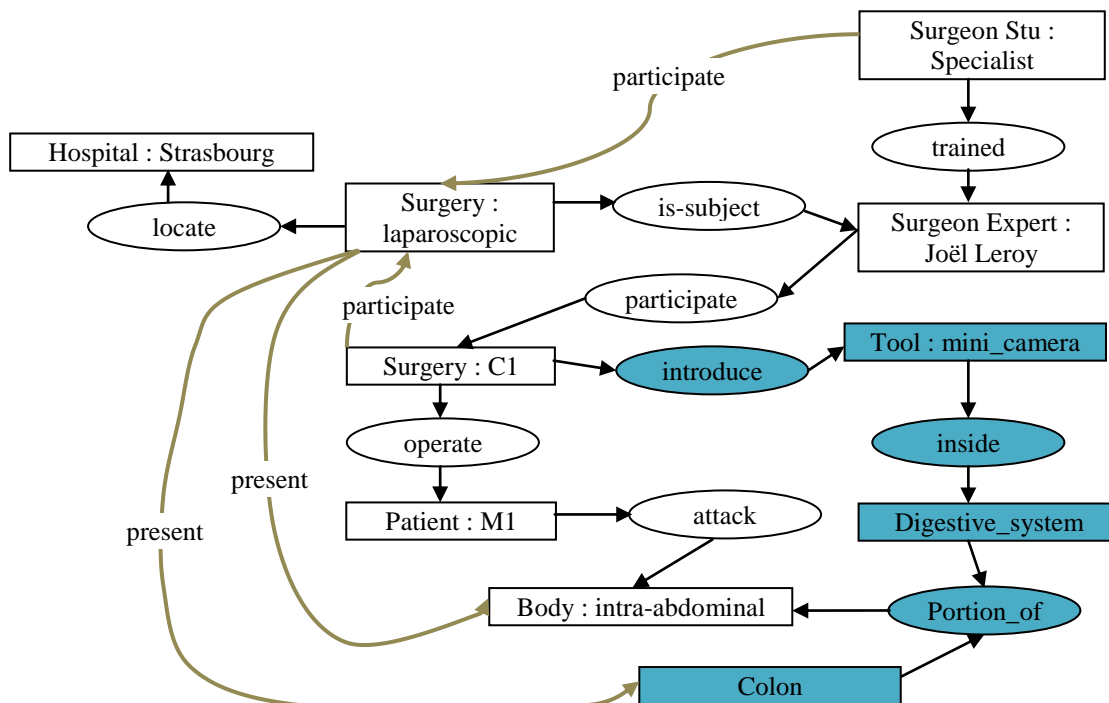


Figure 2. 16: Mise à jour d'Index

III. Représentation sous forme d'un graphe unique

La description structurelle pour la présentation du contenu visioconférence a pour objectif de formuler des descriptions basées sur les caractéristiques physiques d'un contenu visioconférence. Ces caractéristiques sont généralement dépendantes de type média. Pour chaque type média, nous présentons ces informations spécifiques dans un graphe conceptuel.

En utilisant les différentes facettes définies dans la section II du chapitre précédent, une image sera présentée par un graphe conceptuel. Si nous prenons une image de façon aléatoire d'une visioconférence présentant une opération chirurgicale cardiovasculaire pour un type de

rupture d'une dissection d'une artère ayant lieu à « University Hospital, Harrington McLanghlin, heart and vascular Institute », nous présentons cette image par le graphe conceptuel figuré dans la **figure 2.17**.

Pour le contenu audio, nous donnons une représentation abstraite sous forme graphique comme illustré dans la **figure 2.18**.

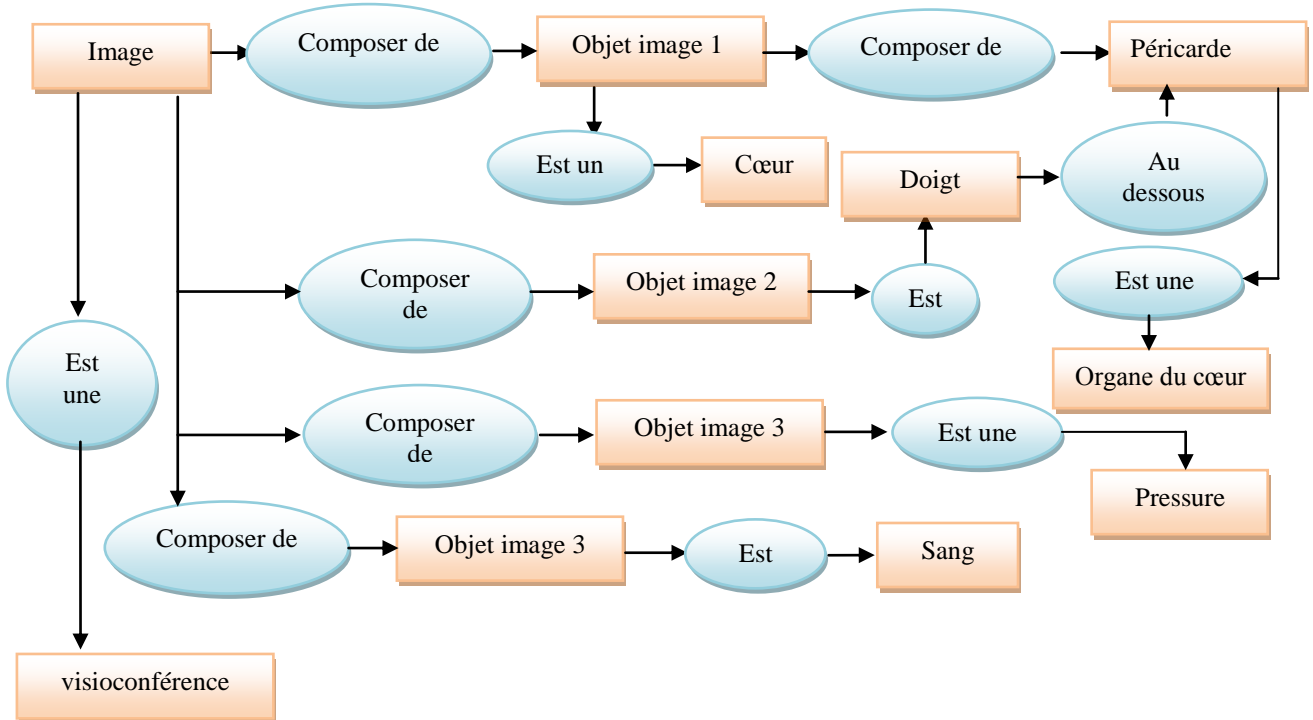


Figure 2. 17: Exemple de représentation formalisme de GC

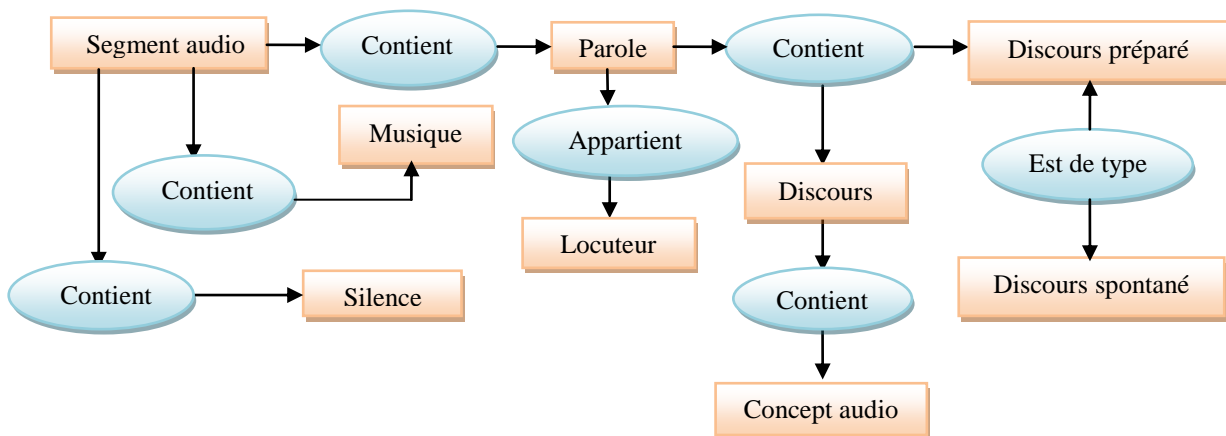


Figure 2. 18 : Description générique du contenu audio

Après l'indexation sémantique, les concepts et les relations sémantiques dégagés seront présentés sous forme de graphe conceptuel.

Le module d'indexation repose sur le triplet {Objet, Relation, Objet} permettant de représenter le contenu des portions visioconférence par des phrases simples exprimées en langage naturel telle que « X parle de Y ».

Les objets visuels composés des entités (des concepts logiques) qui peuvent être reliés à une ou plusieurs portions visioconférence. Un objet visuel peut être défini aussi comme une collection de régions visuelles, qui sont regroupées ensemble sous quelques critères définis par la connaissance du domaine. Ces objets devraient également satisfaire quelques conditions comme la conformité sémantique et la représentation d'un objet réel pour les utilisateurs.

Le « objet image 1 » est un élément d'information rattaché à une ou plusieurs relations. L'interprétation d'une portion visioconférence contenant un ou plusieurs objets, est généralement reliée à un état ou une activité. Il est possible de générer plusieurs interprétations plus ou moins différentes du même concept objet contenu dans la même portion visioconférence.

Vue la diversité du contenu visioconférence, il arrive souvent que pour une même portion visioconférence, nous pouvons associer plusieurs indexations qui se différencient surtout au niveau abstraction et source d'information utilisée (audio, image, texte). La source d'information est un critère de distinction entre les différentes syntaxes de requêtes. En effet, une requête basée sur le flux visuel a la syntaxe suivante « rechercher les portions visioconférence dans lesquelles une personne X apparaît ». Par contre, une requête basée sur le flux audio aura la forme : « rechercher les portions visioconférence dans lesquelles une personne X parle ».

La modélisation du contenu visioconférence avec des concepts tient compte de ces différentes spécificités (audio et visuelle). Il en résulte un ensemble de représentations sous forme de sous-graphes pour chaque type média. Cette distinction n'exclut pas le fait d'avoir une représentation unique du contenu. En effet, si nous prenons l'exemple de la visioconférence chirurgicale détaillée ci-dessus : « L'expert parle de l'opération en cours », cette description peut être interprétée de différentes manières. En se basant sur le contenu audio ou bien par l'intermédiaire de l'information visuelle, nous pouvons par exemple apercevoir une information textuelle à l'écran indiquant qu'il s'agit d'une intervention de l'expert concernant l'opération en cours.

La représentation avec le formalisme des graphes conceptuels est illustrée par l'ensemble des concepts et des relations conceptuelles. Cette présentation peut apparaître dans des graphes séparés bien qu'il s'agisse parfois d'une description de même concept.

Dans l'exemple précédent, il s'agit du même concept « Expert ». Ce concept est interprété différemment : c'est un concept visuel (voir son image) ou bien un concept auditif (entendre sa voix). Pour pouvoir intégrer ces descriptions dans un même schéma, nous proposons d'une part, d'associer une information temporelle à chaque description et d'autre part de spécifier

les types des relations conceptuelles dans les graphes correspondants à chaque flux (audio ou visuel), [Yengui *et al.*, 10].

V. Réseaux des graphes

L'information temporelle permet de structurer les différentes descriptions et de les classer dans le même ordre que leur apparition dans la visioconférence. Les différentes descriptions du document sont regroupées dans un schéma constituant une sorte de réseau de graphes. Nous appelons réseau des graphes, l'ensemble des descriptions sous forme de graphes conceptuels associés aux documents visioconférences. Pour le même exemple d'opération chirurgicale (présenté dans la section II.3.3. du chapitre précédent), un graphe unique sera comme suit, [Yengui *et al.*, 11] :

JOIN (visual graph)

[Io1] → (Vsc) → [Equipe médecins]

[Io2] → (Vsc) → [Salle opératoire]

[Io1] → (Existe dans) → [Io2]

AND (audio graph)

[Ao1] → (is a) → [Expert]

[Ao2] → (is a) → [Opération en cours]

[Ao3] → (is a) → [Equipe médecins]

[Ao4] → (is a) → [Public]

[Ao1] → (parle)

[Ao3] → (parle)

[Ao4] → (parle)

[Ao1] → (parle de) → [Ao2]

[Ao3] → (parle de) → [Ao2]

[Ao4] → (parle de) → [Ao2]

By (Concept node)

[Equipe médecins]

Conclusion

Dans ce chapitre, nous avons présenté les étapes détaillées que nous avons suivies pour l'indexation des visioconférences médicales. Ces étapes sont fondées sur la distance intertextuelle, des mesures statistiques et des ressources sémantiques telles que les ontologies et les thésauri. En première étape, en utilisant des ressources sémantiques, nous comparons la liste des termes à celle des concepts existants dans les ressources. Les concepts trouvés et dénotés par les termes seront ajoutés à une liste de concepts. Ensuite, nous calculons pour chaque concept son poids. Uniquement les concepts pondérés seront retenus. Finalement, et en se basant sur l'hypothèse de Maisonnasse, nous déterminons la liste des relations sémantiques entre les concepts.

Nous signalons que tout au long du processus d'extraction des concepts, les documents seront traduits en anglais vu que nous ne traitons pas l'aspect multilingue dans ce travail. L'extraction des concepts et des relations sémantiques est entièrement automatique et ne nécessite pas l'intervention de l'utilisateur.

Dans le chapitre suivant, nous détaillons le troisième module de notre système : le module de recherche.

Chapitre 3 : Recherche d'information dans une base de visioconférences médicales

Les performances d'un système de recherche d'information peuvent être dégradées en termes de précision du fait de la difficulté pour des utilisateurs à formuler précisément leurs besoins en information. La reformulation ou l'expansion de requêtes constitue une des réponses à ce problème dans le cadre des SRI.

Dans ce chapitre, nous proposons une nouvelle méthode de reformulation de requêtes conceptuelles qui, à partir de la requête initiale de l'utilisateur et d'une ressource externe (RE), cherche un ensemble de concepts maximisant les performances du SRI. Celles-ci sont évaluées, de manière originale, à l'aide d'indicateurs dont une formalisation est proposée.

Puis, nous calculons le score de correspondance entre la requête reformulée et les documents visioconférences médicales pour donner à l'utilisateur le document le plus pertinent

Introduction	136
I.Processus de reformulation	136
I.1. Phase de capture du contexte statique	138
I.2. Phase de reformulation	138
II.Étape d'appariement document-requête	142
II.1. Appariement de graphes	143
II.1.1. Projection de graphes	143
II.1.2. Calcul de similarité	144
II.2. Calcul de pertinence.....	146
Conclusion	147

Introduction

L'utilisateur trouve souvent des difficultés pour traduire son besoin exact en information. Par conséquent, parmi les documents qu'ils sont retournés par l'outil de recherche, certains documents ne satisfont pas son besoin.

Pour résoudre ce problème et pour rendre à l'utilisateur uniquement les documents les plus pertinents, une amélioration de la requête initiale proposée est nécessaire. Cette amélioration est faite par un processus de reformulation de la requête initiale. Ce processus prend en compte, en premier lieu, le PU et en deuxième lieu, il prend la génération d'une nouvelle requête en utilisant les REs.

La requête reformulée obtenue sera envoyée vers le processus d'appariement Documents-Requête. Durant ce processus, nous calculons les mesures de similarité entre la requête reformulée et les différentes visioconférences de la base de données. Cette comparaison est faite par la projection du graphe conceptuel de la requête reformulée sur tous les graphes conceptuels des documents obtenus après le module d'indexation.

Nous détaillons dans ce qui suit le processus de reformulation des requêtes et le processus d'appariement Documents-Requête.

I. Processus de reformulation

C'est un processus évolutif et interactif. Son principe est d'utiliser la requête initiale pour amorcer la recherche, pour modifier celle-ci à partir des jugements de pertinence et/ou de non-pertinence de l'utilisateur dans le but de ré-pondérer les termes de la requête initiale, ou y ajouter (ou supprimer) d'autres termes contenus dans les documents pertinents (ou non pertinents). La nouvelle requête obtenue permet de corriger la direction de la recherche pour le sens des documents pertinents.

Dans notre travail, nous proposons une reformulation des requêtes guidées par une RE. Il s'agit de réécrire la requête de l'utilisateur en prenant en compte les relations de synonymie et les relations de méronymie présentes dans les ressources externes. Ainsi, la requête est enrichie par les termes sémantiquement proches des termes d'origine de la requête. Ces termes proches sémantiquement sont issus des ontologies de domaine afin de rechercher des documents dans un corpus médical.

De même, nous nous basons sur le PU pour la reformulation des requêtes. Nous utilisons un mécanisme progressif pour catégoriser les utilisateurs en construisant une base d'éléments contextuels statiques servant à reformuler la requête initiale afin de produire une nouvelle requête qui reflète au mieux le besoin de l'utilisateur, [Yengui *et al.*, 14].

La structure de l'étape de reformulation, les différentes tâches qui la composent ainsi que les interactions entre elles sont illustrées par la **figure 2.19**.

La requête reformulée est appariée aux index des documents dont les index possédant une « similarité » suffisante à la requête sont considérés être pertinents. Ci-dessous nous présentons un exemple de modification des requêtes.

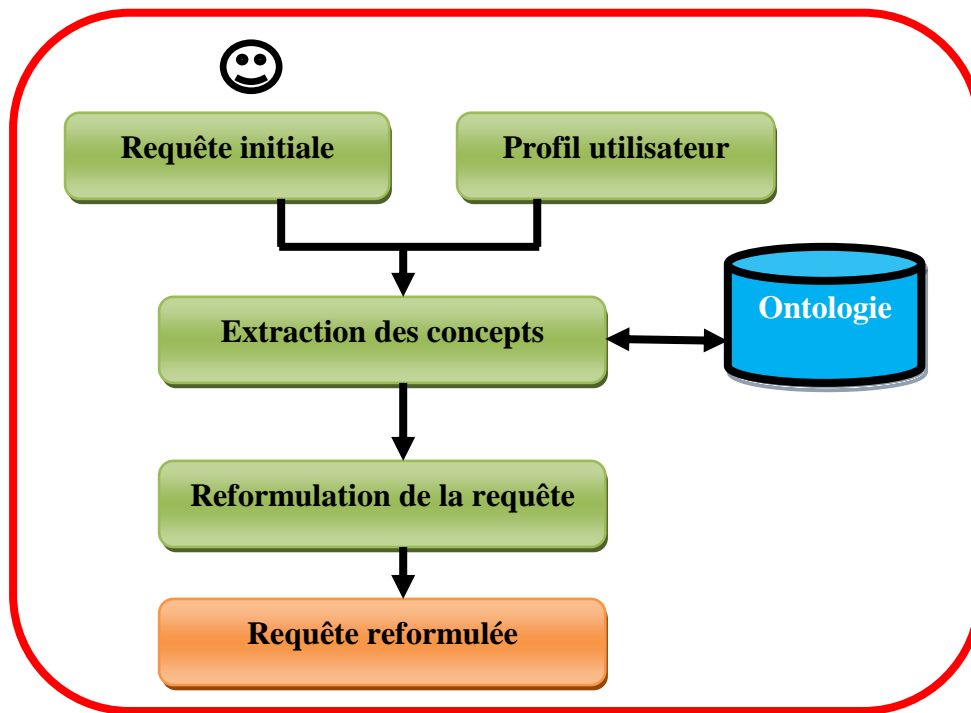


Figure 2. 19: Reformulation de la requête utilisateur

Tableau 2. 4: Exemples de modification de requête

	Q = migraine (requête initiale)	
Type 1	Q' = Q + faqs (affinement de la requête par ajout d'un terme)	Modification de la requête par rapport à un concept du corpus.
Type 2	Q' = migraine ou céphalée (élargissement de la requête à un domaine plus large) Q' = migraine + aura (affinement de la requête par ajout d'un terme)	Modification de la requête par rapport au domaine.

La reformulation des requêtes se déroule en trois phases successives : **(i)** capture du contexte statique, **(ii)** identification des concepts qui appartiennent le besoin de l'utilisateur, **(iii)** composition de ces termes dans le but de formuler la requête. Notons en général, les requêtes telles qu'elles sont formulées par les utilisateurs suivant les cas ciblent : **(i)** des concepts ou des termes relatifs au domaine de la requête : migraine, « aura migraineuse », céphalées ; **(ii)** des éléments présents dans certaines parties d'une visioconférence (type de média).

Afin de bien formuler la requête, nous proposons de faire appel à l'ontologie et/ou le thésaurus. Il s'agit de recourir à certaines relations sémantiques qui jouent un rôle significatif dans l'organisation des connaissances (relation causale, définition, ..).

I.1. Phase de capture du contexte statique

Cette phase sert à identifier un utilisateur à travers une série d'information pour le catégoriser. L'utilisateur définit son contexte statique lors de sa première utilisation du système. Ce contexte est composé par quatre catégories, [Yengui *et al.*, 14] :

- Paramètres de connexion : login, mot de passe
- Caractéristiques personnelles : nom, prénom, pays, ...
- Intérêt et préférences : domaine, spécialité, ...
- Compétences : Profession, niveau d'étude, ...

Cette phase est présentée par la **figure 2.20**.

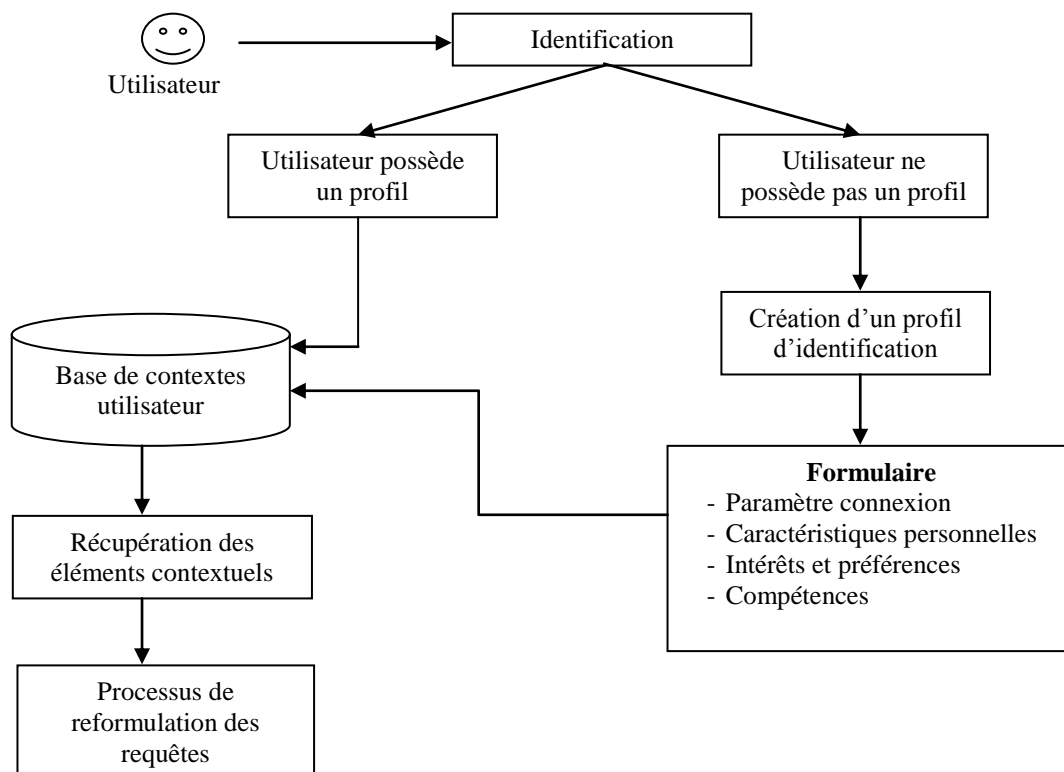


Figure 2. 20 : Phase de récupération du contexte statique des utilisateurs

I.2. Phase de reformulation

Cette phase a pour objectif de produire une nouvelle requête à partir de la requête initialement formulée par l'utilisateur et cela en rajoutant des termes issus de son contexte de recherche actuelle. L'utilisateur donne sa requête en utilisant ces propres termes. Le système procède à l'extraction de l'ensemble des termes à rajouter afin de produire une nouvelle requête. En premier lieu, ces termes sont extraits de la base des contextes utilisateurs. Puis nous récupérons les concepts spécifiques à partir de l'ontologie. Dans le cas où le concept *C1* est présenté comme terme dans la requête initiale, il sera expansé par le concept *C2* récupéré à partir de la ressource (UMLS et/ou OntoMénélas).

L'élargissement des concepts est basé sur la présence d'un lien sémantique entre $C1$ et $C2$ (synonyme ou concept plus général).

Cette phase se déroule selon deux étapes agissant sur différents aspects de la requête initiale, ces étapes sont :

- **Extraction des concepts** : nous récupérons les termes de la requête présents dans la ressource. Puis, nous parcourons cette dernière en utilisant les termes comme point d'entrée afin d'extraire les concepts liés directement à chaque terme, [Yengui *et al.*, 14].

Les concepts sont récupérés de la manière suivante :

- Si le concept ($C1$) est présent comme terme dans la requête initiale, il sera expansé par le concept ($C2$) récupéré à partir de la RE.
- Le choix de ($C2$) est fait suite à un parcours du fichier XML décrivant la RE.
- Nous cherchons un lien sémantique entre ($C1$) et ($C2$). Ce lien peut être de type « Synonyme », « Concept plus général », ...

Dans ce cas, le rôle de l'utilisateur est passif car il n'intervient pas dans la sélection des concepts.

Le temps de calcul est trop élevé pour une ontologie de grande taille. Dans le cadre d'une approche heuristique, il est raisonnable de ne tester que des concepts au voisinage sémantique des concepts indexant les documents. Si $\prod(C_X, C_Y)$ est une mesure de proximité sémantique entre C_X et C_Y deux concepts de Θ , nous construisons $C(\Theta)$ de la manière suivant :

$$c(\theta) = \{C_i \in \theta | \exists C_j \in C_{(D)}, \prod(C_j, C_i) > \varepsilon\} \quad (2.6)$$

$C(\Theta)$ est construite en identifiant pour chaque concept, les concepts de son voisinage qui doivent être considérés, [Sy *et al.*, 12].

- **Reformulation de requête** : l'intégration dans la phase de la reformulation de requête consiste à argumenter la requête initiale par les concepts issus de l'étape d'extraction des concepts et des termes issus du PU. La nouvelle requête se transmet par la suite à l'étape d'appariement (section II). Cette requête sera présentée sous la forme d'un graphe conceptuel, [Yengui *et al.*, 14].

Le processus de reformulation est résumé par la **figure 2.21** et l'algorithme suivant :

Algorithme Reformulation
<p>Entrée :</p> <p>Q : Requête utilisateur</p> <p>O : Ontologie</p> <p>P : Profil utilisateur</p>
<p>Sortie :</p> <p>Q' : Requête reformulée</p>
<p>Variables :</p>

<p>Lc : Liste des concepts Lt : Liste des termes</p>
<p>Début Prétraitement de la requête ; /* Extraction des termes*/ Si (existe-déjà (Lt) = vrai) alors Lc ← Lt ; Sinon Lc ← Extraction (Lt) ; Fin si Q' ← Q + Lc Fin</p>

Algorithme 9: Reformulation d'une requête

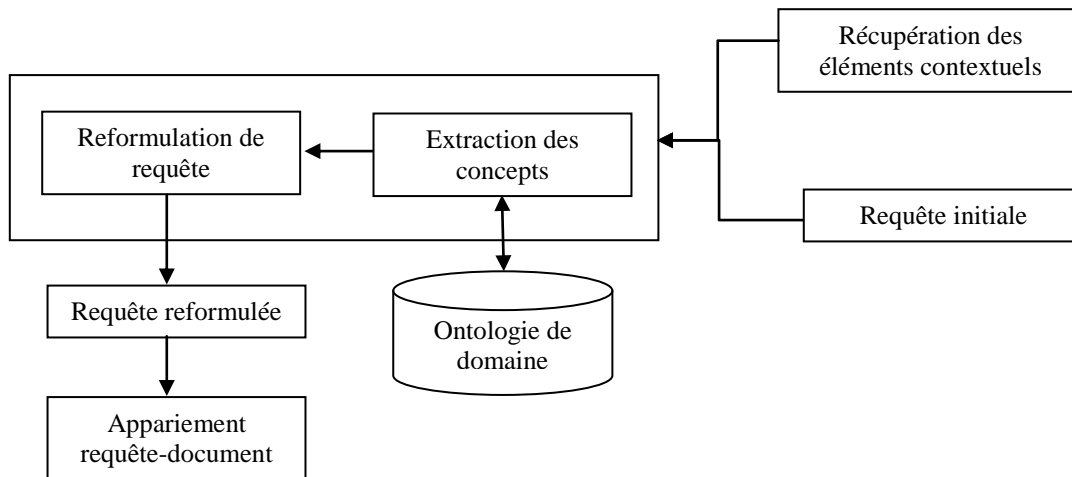


Figure 2. 21 : Architecture du processus de reformulation

Comme il est décrit dans l'algorithme de reformulation d'une requête, notre approche fait appel à un algorithme d'extraction qui permet d'identifier les concepts à chaque terme. Cet algorithme est le suivant :

<p>Algorithme Extraction</p>
<p>Entrée : Lt : Liste des termes Lco : Liste des concepts de l'ontologie Iu : Intérêt de l'utilisateur Cu : Compétence de l'utilisateur</p>
<p>Sortie : Lc : Liste des concepts</p>
<p>Variables :</p>

<p>t : Terme c : Concept</p>
<p>Début</p> <p>Pour chaque terme t faire</p> <p> Identifier les concepts associés à t</p> <p> Selon Iu faire</p> <p> Cas1 : Iu est « non médecine » :</p> <p> Si t appartient à Lco alors</p> <p> Ajouter (c(t)) à Lc /* c est le concept identifié */</p> <p> Fin si</p> <p> Cas2 : Iu est « médecine » :</p> <p> Selon Cu faire</p> <p> Cas1 : Cu est « expert » ou « médecin » :</p> <p> Si t appartient à Lco alors</p> <p> Ajouter (c) à Lc /* c est le concept identifié */</p> <p> Fin si</p> <p> Cas2 : Cu est « étudiant » :</p> <p> Si t appartient à Lco alors</p> <p> Ajouter (c(t)) à Lc /* c est le concept identifié */</p> <p> Fin si</p> <p> Fin selon</p> <p> Fin selon</p> <p> Fin pour</p> <p>Fin</p>

Algorithme 10: Extraction des concepts

Exemple :

En prenant l'exemple de la requête « les maladies du cœur causent le rythme rapide du cœur » et en utilisant l'algorithme de reformulation proposé ci-dessous, nous trouvons :

- **Entrée :**
 - Q : « les maladie du cœur causent le rythme rapide du cœur »
 - OntoMénélas : Ontologie médicale
 - Iu : « Médecine »
 - Cu : « Expert »
- **Sortie**
 - Q' : nouvelle requête (détaillé dans la suite)
- **Variables**

- Lc : liste des concepts (détaillé dans la suite)
- Lt : Liste des termes (détaillé dans la suite)

En se basant sur l'algorithme d'Extraction des termes simples mentionnés dans le premier chapitre, nous extrayons, la liste des termes simples suivants :

Lts = {maladie ; cœur ; rythme ; rapide}.

Puis en utilisant l'algorithme d'Extraction des termes composés mentionnés dans le même chapitre et à partir de la liste des termes simples, nous obtenons la liste des termes composés suivante :

Ltc = {maladie_cœur ; rythme rapide}.

La combinaison de ces deux listes est considérée comme la liste des termes à utiliser dans l'algorithme de reformulation. D'où

Lt = {maladie ; cœur ; maladie_cœur ; rythme ; rapide ; rythme_rapide}.

Ensuite, en appelant l'algorithme d'extraction, nous identifions les concepts en les associant des termes. Cette identification est présentée dans le tableau suivant :

Tableau 2. 5: Extraction des concepts

Terme	Concept
Maladie_cœur	Cardiovasculaire
Maladie	Maladie
Cœur	Cœur
Rythme_rapide	Tachycardie
Rythme	Rythme
Rapide	Rapide

D'où la requête reformulée sera :

Q' = « Maladie + Cœur + Cardiovasculaire + Rythme + Rapide + Tachycardie »

Cette requête reformulée sera envoyée par la suite à l'étape d'appariement document-requête. Cette étape se base essentiellement sur une fonction de correspondance pour satisfaire le besoin de l'utilisateur.

II. Étape d'appariement document-requête

Le traitement de la requête exploite un algorithme d'appariement de graphes (graph matching) afin d'extraire les visioconférences qui correspondent à la sélection opérée par la requête. La requête obtenue après l'étape de reformulation est présentée à l'aide d'un graphe conceptuel. Par exemple pour la même requête « les maladies du cœur causent le rythme rapide du cœur » (utilisé dans la section I.2 de ce chapitre), et après la reformulation de cette

requête, nous avons obtenu la liste des concepts suivants « cardiovasculaire, tachycardie » et la relation « cause ». Cette requête peut s'écrire :

« [Maladie : # Cardiovasculaire] → (cause) → [Maladie : # Tachycardie] ».

La recherche des visioconférences correspondantes à une requête se fait en deux étapes, [Yengui *et al.*, 14] :

- Appariement de graphes : cette étape se fait, elle-même, en deux étapes. En premier lieu, nous effectuons une projection du graphe conceptuel de la requête sur les graphes conceptuels des documents. Dans le cas où cette tâche ne rend pas de résultat, nous utilisons, en second lieu, une fonction de similarité entre les concepts.
- Calcul de pertinence des documents trouvés.

II.1. Appariement de graphes

II.1.1. Projection de graphes

Le formalisme de graphes conceptuels permet d'interroger et de chercher des sous-graphes dans un graphe en utilisant la projection, [Sowa, 84]. Cette dernière prend en compte les concepts et les relations. La projection d'un graphe requête G_{Ri} sur un graphe conceptuel G_{Ai} notée $\Pi G_{Ai}(G_{Ri})$, conclue à l'existence d'un sous-graphe de G_{Ai} qui est spécifique au graphe G_{Ri} . De manière informelle, un graphe conceptuel G_D est une projection d'un graphe G_R si chaque concept de G_R a un concept spécifique dans G_D . Si une telle projection existe, alors il a été prouvé que le document répond à la requête (cf. **Figure 2.22**).

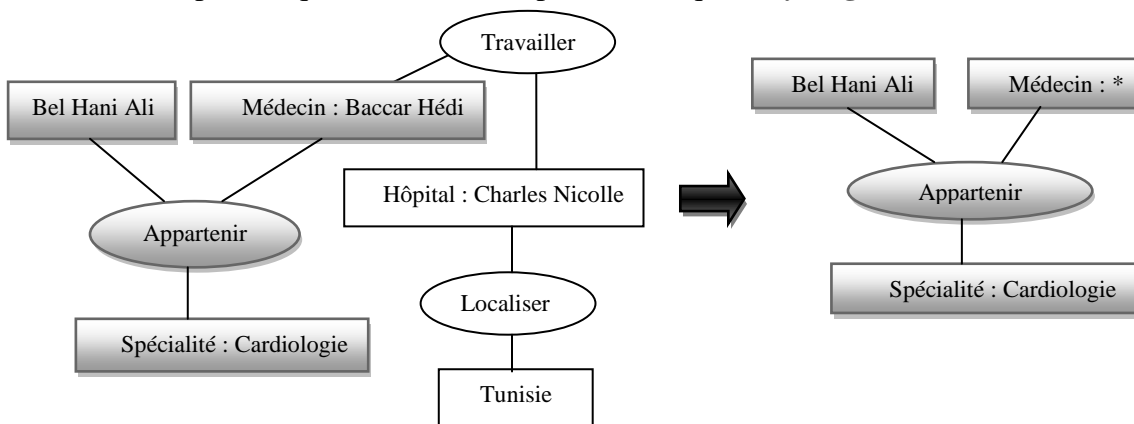


Figure 2. 22: Opération de projection des GCs

Dans [Ounis *et al.*, 98], il a été montré qu'un algorithme de recherche correspondant à l'opérateur de projection sur les graphes conceptuels peut être implémenté de manière très efficace en termes de complexité. Nous proposons d'utiliser cet algorithme. Nous quantifions l'appariement entre un graphe requête G_R et un graphe d'indexation G_D en combinant un appariement sur les concepts et un appariement sur les arches. Nous appelons arche un triplet (concept, relation, concept) qui lie trois nœuds – deux concepts et une relation – dans un graphe conceptuel.

D'où

$$F(G_R, G_D) = \sum \left\{ ft(C).fid(C) \mid C \in \text{Concepts de } \prod G_D(G_R) \right\} + \sum \left\{ ft(A).fid(A) \mid A \in \text{Arches de } \prod G_D(G_R) \right\} \quad (2.7)$$

Dans la formule (2.7), les fréquences de termes ft et les fréquences inverses de documents fid sont calculées comme suit :

- La fréquence du terme ft associée à un concept C dans un graphe conceptuel G est définie comme le nombre de concepts de G qui sont spécifiques au concept C .
- La fréquence inverse de document fid associée à un concept C est basée sur les documents qui sont décrits par C ou par un concept spécifique à C . Nous utilisons une formule inspirée de [Salton *et al.*, 98] :

$$fid(C) = \log_{10}(1 + D|d(C)) \quad (2.8)$$

Où :

- D : le corpus des documents
- d : les documents correspond à C .

Pour une arche, le principe est similaire aux concepts.

Etant donné qu'une indexation peut être un graphe conceptuel, ou un ensemble de graphes conceptuels non connexes, nous définissons l'appariement entre un graphe requête G_R et un ensemble S de graphe d'indexation G_D , comme le maximum des appariements entre G_R et chacun des graphes G_D de S :

$$M(G_R, S) = \max_{G_D \in S} (F(G_R, G_D)) \quad (2.9)$$

II.1.2. Calcul de similarité

Dans le cas où la projection du graphe requête sur les graphes d'indexation ne donne aucun résultat, nous utilisons une mesure de similarité entre les graphes conceptuels associés à la requête et ceux associés aux différents documents du corpus.

Dans notre système, nous prenons en compte les différents termes constituant un concept. La similarité entre deux concepts est assimilée au nombre de termes communs entre eux. Cette fonction est appelée *Sim*. Plus que deux concepts ont des termes communs, plus qu'ils sont proches l'un de l'autre.

En premier lieu, nous calculons la proximité d'un concept donné. Cette proximité est mesurée par le rapport entre les termes communs (de deux concepts) par rapport à tous les termes du concept.

Nous proposons :

- C_R : concept trouvé dans la requête ;
- $Terme(C_R)$: l'ensemble des termes existants dans C_R ;
 $Terme(C_R) = \{A, B, D, H, M\}$
- C_D : concept trouvé dans la visioconférence ;

- $Terme(C_D)$: l'ensemble des termes existants dans C_D ;
 $Terme(C_D) = \{A, B, D, G, L\}$

Nous appelons les termes communs entre les deux concepts par *Ancêtre*.

- $Ancêtre(C_R, C_D) = Terme(C_R) \cap Terme(C_D)$
 $= \{A, B, D, H, M\} \cap \{A, B, D, G, L\}$
 $= \{A, B, D\}$

La proximité d'un concept C_R par rapport aux $Ancêtre(C_R, C_D)$ est exprimée par :

$$Proximité(C_R) = \frac{Cardinal (Ancêtre (C_R, C_D))}{Cardinal (Terme (C_R))} \quad (2.10)$$

La proximité d'un concept C_D est calculée avec le même principe.

$$Proximité(C_D) = \frac{Cardinal (Ancêtre (C_R, C_D))}{Cardinal (Terme (C_D))} \quad (2.11)$$

De même, nous proposons la formule de similarité sémantique entre les concepts :

$$Sim(C_R, C_D) = Proximité(C_R) * Proximité(C_D) \quad (2.12)$$

Prenons l'exemple cité au-dessus :

- $Terme(C_R) = \{A, B, D, H, M\}$
- $Terme(CD) = \{A, B, D, G, L\}$
- $Ancêtre(C_R, C_D) = \{A, B, D\}$
- $Proximité(C_R) = \frac{Cardinal \{A,B,D\}}{Cardinal \{A,B,D,H,M\}} = 3/5 = 0,6$
- $Proximité(C_D) = \frac{Cardinal \{A,B,D\}}{Cardinal \{A,B,D,G,L\}} = 3/5 = 0,6$
- $Sim(C_R, C_D) = 0,6 * 0,6 = 0,36$

Nous calculons la similarité de chaque concept de la requête avec ceux des documents.

Cette similarité permet de favoriser les visioconférences comprenant de nombreux concepts de la requête, [Yengui *et al.*, 14].

Algorithme Appariement Document-Requête
<p>Entrée :</p> <p>R : Requête</p> <p>V : Corpus de visioconférence</p>
<p>Sortie :</p> <p>Sim(R, D) : valeur de similarité entre R et D</p>
<p>Variables :</p> <p>C_R : Concept de la requête</p>

<p>C_D : Concept de document D : Document de la visioconférence Sim_i : similarité entre CR et CD</p>
<p>Début /* Calcul de similarité entre requête et document*/ $Sim(R, D) \leftarrow 1$ Pour chaque Document D du Corpus V faire Pour chaque Concept C_D de D faire $Sim_i \leftarrow$ Calculer la similarité entre C_D et C_R $Sim(R,D) \leftarrow Sim(R, D) * Sim_i$ Fin pour Fin pour Si $Sim(R, D) > \text{seuil_similarité}$ alors Garder (D) Finsi Fin</p>

Algorithme 11: Appariement Document-Requête

II.2. Calcul de pertinence

Dans la section précédente, nous avons calculé la similarité entre les graphes conceptuels requêtes et les graphes conceptuels documents visioconférence et nous avons affecté à chaque document visioconférence un score de similarité. Nous pouvons obtenir plusieurs documents correspondants à la requête. Il est nécessaire de les ordonner.

Uniquement les documents qui ont un score de similarité non nul, seront triés par ordre décroissant. Le module de recherche retournera soit les concepts les plus pertinents par document, soit l'ensemble de tous les concepts pertinents triés par ordre décroissant groupés par document, et les documents seront triés à leur tour par leur pertinence par rapport à l'ensemble des concepts de la requête.

Le score de pertinence est calculé par la formule suivante :

$$Score(R, D) = \sum \text{Pondération } C_D * Sim(C_R, C_D) \quad (2.13)$$

Où :

- R est une requête ;
- D est la visioconférence ;
- Pondération C_D : la pondération du concept dans la visioconférence.

Ce score permet de favoriser les visioconférences comprenant de nombreux concepts de la requête, [Yengui *et al.*, 14].

Algorithme Pertinence Document-Requête
Entrée : R : Requête V : Corpus de visioconférence
Sortie : S : Score de pertinence d'un document D / Requête
Variables : C_R : Concept de la requête C_D : Concept de document D : Document de la visioconférence Sim(R, D) : Valeur de similarité entre R et D Sim_i : Similarité entre C_R et C_D
Début Score(R, D) \leftarrow 1 /* Calcul de score de pertinence d'un document */ Pour chaque Document D du Corpus V faire Pour chaque Concept C_D de D faire Score(R, D) \leftarrow Pondération C_D * Sim_i Fin pour Fin pour Si Score(R, D) > seuil_pertinence alors Garder Score(R, D) Finsi Fin

Algorithme 12: Pertinence Document-Requête

Conclusion

Le module de recherche se compose essentiellement de deux étapes. La première est la reformulation de requête initiale en exploitant les liens sémantiques entre les concepts en utilisant une ontologie de domaine. C'est une étape évolutive et interactive. Elle utilise la requête initiale pour amorcer la recherche, dans le but de ré-pondérer les termes de la requête initiale, ou y ajoute (ou on lui supprime) d'autres termes. La nouvelle requête obtenue permet de corriger la direction de la recherche pour le sens des documents pertinents.

Cette reformulation est guidée par une RE pour prendre en compte des relations de synonymie et ses relations de méronymie qui peuvent être présentées sans oublier l'ajout des termes sémantiquement proches des termes d'origine de la requête.

La deuxième étape est l'appariement requête-document. Elle se base sur une fonction de correspondance qui permet de comparer la représentation de la requête à celle de chaque

document. Les documents retenus seront classés en utilisant la fonction de similarité et présentés à l'utilisateur par le triplet (titre et concepts trouvés).

Les travaux présentés dans cette partie seront réalisés et évalués dans la troisième.

Partie 3

*Réalisation : Expérimentation et
évaluation*

Chapitre 1 : Implémentation

Ce chapitre décrit l'implémentation du système OSSVIRI qui présente un SRI pour les documents visioconférences médicales. Ce système est composé de trois modules : Analyse, Indexation et Recherche.

Introduction.....	151
I. Récapitulatif des apports méthodologiques de notre thèse.....	151
II. Expérimentation	152
II.1. Environnement de développement.....	152
II.1.1.Choix du Système d'exploitation.....	152
II.1.2. Choix de l'outil d'exploitation de l'ontologie.....	152
II.1.3. Choix du Langage de programmation.....	152
II.1.5. Choix du SGBD	153
II.2. Description du système OSSVIRI.....	153
II.2.1. Menu Fichier	154
II.2.2. Module d'Analyse.....	155
II.2.2. Module d'Indexation.....	157
II.2.4. Module de Recherche.....	160
Conclusion	162

Introduction

Après avoir effectué la conception d'un SRI du document audiovisuel en exploitant deux ressources externes, nous allons à présent entamer la réalisation de ce SRI que nous avons baptisé « OSSVIRI » (**O**ntology-based **S**ystem for **S**emantic **V**isio-conference **I**nformation **R**etrieval and **I**ndexation).

Nous présenterons alors, dans la première partie de ce chapitre, l'ensemble des outils de développement utilisés. Dans la seconde partie, nous détaillerons le processus d'implémentation de l'OSSVIRI, ainsi que les principales interfaces qui le composent à travers des fenêtres de capture.

I. Récapitulatif des apports méthodologiques de notre thèse

Le principe de notre thèse est d'élaborer un SRI sémantique des documents audiovisuels. Nous avons constaté qu'indexer est une tâche essentiellement interprétative d'analyse et de reformulation qui demande beaucoup de savoir et de savoir-faire. De plus, le résultat de recherche -dans la base d'index- doit tenir compte de tous les besoins « imprécis » des utilisateurs. Le problème existant est que la continuité sémantique doit être conservée et donne des réponses exactes aux besoins des utilisateurs.

L'état de l'art recommandé pour contrôler cette continuité sémantique est le recours à des langages documentaires. Ces derniers permettent de contrôler les expressions qui servent d'index, mais aussi de fournir des référentiels qui guident l'interprétation et l'indexation. Pour améliorer cette indexation, nous avons pris en considération les points suivants :

- **L'expressivité autorisée pour les index** : les contenus documentaires sont complexes, spécialement pour les documents audiovisuels. D'où la nécessité d'introduire un vocabulaire conceptuel riche et adapté à l'application envisagée. De plus, les index créés doivent lier explicitement les éléments conceptuels à l'aide des notions relationnelles spécifiques au domaine médical.
- **Le contrôle** : la définition rigoureuse du référentiel sémantique du langage impose deux efforts de spécification. Le premier, de type informel et manuel, est dirigé vers les besoins des utilisateurs pour les interpréter de façon suffisamment précise. Et le deuxième, de type formel, permet de vérifier la validité sémantique des index produits par notre système.
- **La manipulation des descriptions** : le système peut prendre en charge les tâches les plus simples de réinterprétation et de complétion que les utilisateurs accomplissent lorsqu'ils sont confrontés aux index lors de leurs recherches.

Nous avons vu tout au long du second chapitre comment les fonctionnalités offertes par les systèmes ontologiques répondaient à nos besoins en termes de conception référentiels, sémantiques et formels : création de hiérarchies à la signification précise, définitions formelles des concepts, règles de composition relationnelles... Nous devons donc exploiter ces significations pour compléter les informations fournies par les index et mieux répondre aux requêtes posées par les utilisateurs.

Mais parfois, nous trouvons des concepts trouvés dans un DAV qui n'existent pas dans l'ontologie. Ces concepts doivent être ajoutés dans l'ontologie utilisée. Tous les concepts retenus dans la base d'index seront par la suite représentés sous forme de graphes conceptuels.

Un tel cadre méthodologique permet, dans le contexte de l'indexation à base de connaissance, une utilisation des ontologies qui se fassent de façon mieux adaptée aux besoins que nous avons fixés. Aussi, il est le fruit d'expérimentation à différentes échelles pour le domaine médical en utilisant un corpus de visioconférences médicales. Ces expérimentations ont permis d'observer des problèmes concrets et de tester les outils que nous avons réalisés. Nous allons à présent détailler les plus significatives entre elles.

II. Expérimentation

Nous avons fait le choix de nous concentrer sur le domaine médecine. Nous avons extrait une centaine de documents de type visioconférences médicales. Il s'agit donc d'une collection homogène, ce qui facilite la recherche d'une RE thématique adaptée. De fait, comme le sujet de médecine a attiré l'attention de plusieurs chercheurs, ces dernières années, une grande quantité de ressources sont disponibles dans ce domaine.

II.1. Environnement de développement

Avant de commencer l'implémentation de l'architecture conceptuelle de notre SRI, nous allons tout d'abord spécifier les outils utilisés qui nous ont semblés être un bon choix de part les avantages qu'ils offrent.

II.1.1. Choix du Système d'exploitation

Notre application a été développée sous le système d'exploitation Windows 7, mais comme elle est développée en langage java, elle peut être intégrée dans n'importe quel autre système d'exploitation supportant la machine virtuelle java (Windows 98/00, Linux, ...).

II.1.2. Choix de l'outil d'exploitation de l'ontologie

Les ressources externes utilisées sont implémentées en langage OWL, or les fichiers OWL sont inexploitable en état brut car leur structure est très complexe. Donc pour pouvoir l'exploiter, il nous a fallu un « traducteur » capable de traduire les balises et la sémantique véhiculée par le fichier OWL en objet manipulable par des programmes. L'outil disponible qu'on a pu avoir est L'API JENA. Cet outil est développé par une équipe de la firme **H**ewlett **P**ackard (HP) dans le cadre du Projet HP « *Labs Semantic Web Programme* » qui a pour but de réaliser un outil d'exploitation des fichiers OWL. JENA est développé entièrement en Java, elle donne aux programmes la possibilité d'exploiter le contenu des fichiers RDF et OWL (extraction du contenu sémantique de ces derniers).

II.1.3. Choix du Langage de programmation

Pour le choix de programmation de notre système nous avons opté pour le langage JAVA et cela pour de nombreuses raisons :

- JAVA est un langage orienté objet simple, qui réduit le risque des erreurs d'incohérences.

- Il est indépendant de toute plateforme, il est possible d'exécuter des programmes JAVA sur tous les environnements qui possèdent un **J**ava **V**irtual **M**achine (JVM).
- Il est doté d'une riche bibliothèque de classes, comprenant la gestion des interfaces graphiques (fenêtres, menus, graphismes, boîtes de dialogue, contrôles), la programmation multithread (multitâche), la gestion des exceptions.
- Il permet d'accéder d'une manière simple aux fichiers et aux réseaux (notamment Internet).
- Il permet un accès aux bases de données simplifié soit à travers la passerelle JDBC-ODBC ou à travers un pilote JDBC spécifique au SGBD.
- Il est caractérisé aussi par la réutilisation de son code ainsi que la simplicité de sa mise en œuvre.
- Il existe une **I**nterface de **p**rogrammation d'**a**pplications (API) JAVA fournie avec l'éditeur d'ontologies Protégé-OWL ce qui permet d'accéder à l'ontologie à partir de notre application.
- Il est compatible avec L'API JENA, ce qui nous permet la manipulation, le parcours et la modélisation des documents OWL.

II.1.5. Choix du SGBD

Le Système de Gestion de Bases de Données que nous avons choisi pour implémenter notre base d'Index est Oracle Data base 10g Express édition. Oracle est un **S**ystème de **G**estion de **B**ase de **D**onnées **R**elationnelles (SGBDR) fonctionnant sur diverses plates-formes matérielles sous différents systèmes d'exploitation.

II.2. Description du système OSSVIRI

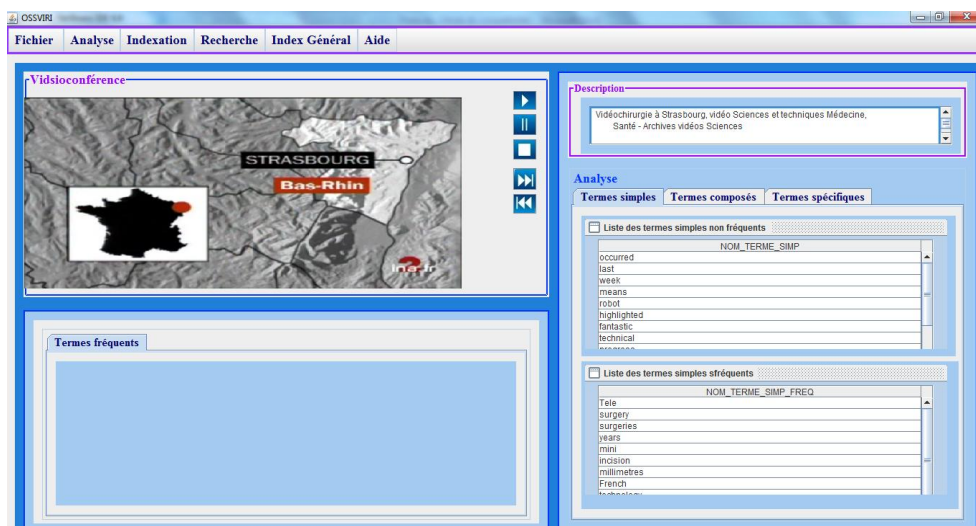


Figure 3. 1 : Interface Module Analyse d'OSSVIRI

L'interface homme/machine représente l'élément clé dans l'utilisation de tout système informatique. Les interfaces de notre SRI sont conçues de manière à être simples, naturelles,

et de compréhension et d'utilisation faciles. La maquette écran suivante représente l'interface principale d'OSSVIRI.

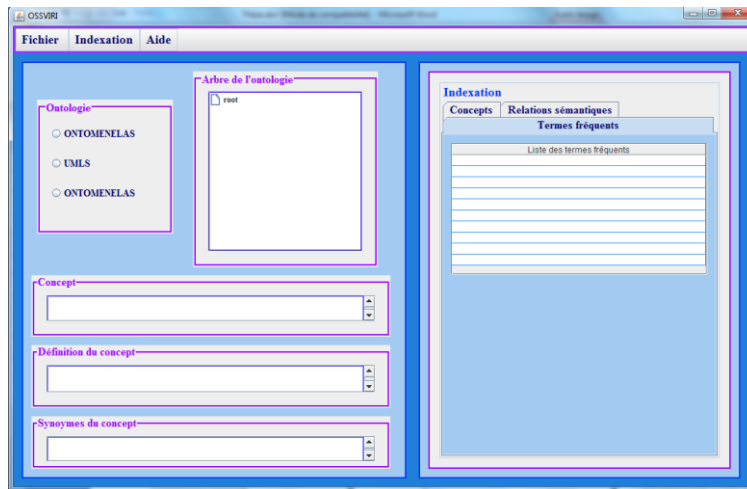


Figure 3. 2 : Interface Module Indexation d'OSSVIRI

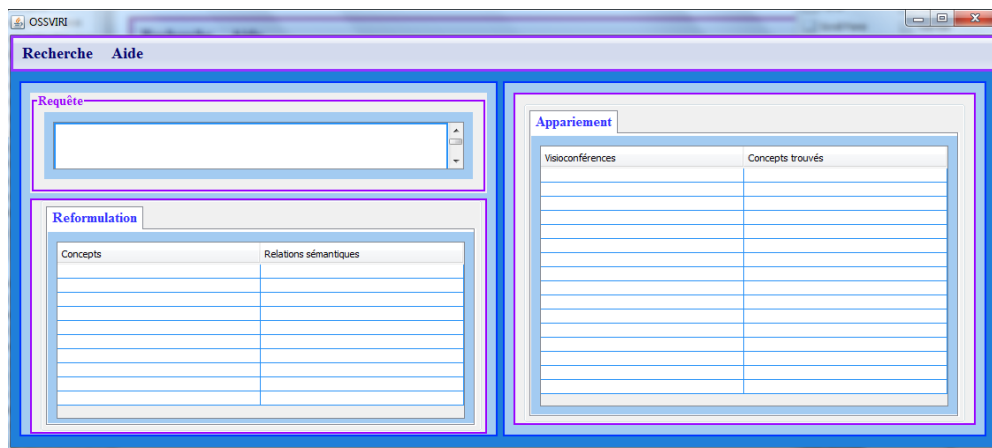


Figure 3. 3 : Interface Module Recherche d'OSSVIRI

II.2.1. Menu Fichier

Le menu Fichier contient deux commandes : « **Importer** » et « **Quitter** ». La commande « **Importer** » permet d'importer une visioconférence en donnant son chemin d'accès.

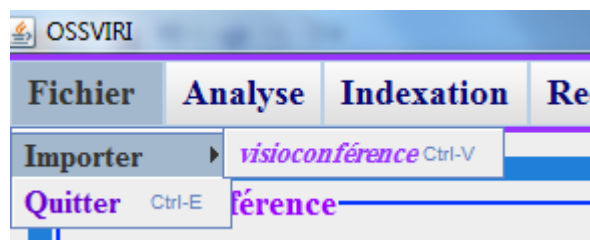


Figure 3. 4 : Menu Fichier

II.2.2. Module d'Analyse

Le module d'analyse d'OSSVIRI est composé de trois étapes :

- 1-Extraction des termes simples en supprimant les mots vides et les mots dont leur fréquence est inférieure à un seuil donné et par la suite l'extraction des termes simples fréquents (cf. **Algorithme 1**).
- 2-Extraction des termes composés en s'aidant des calculs de l'information mutuelle adaptée IMA et par la suite l'extraction des termes simples fréquents (cf. **Algorithme 2**).
- 3-Extraction des termes spécifiques.

Le menu **Analyse** contient trois sous commandes « **termes simples** », « **termes composés** » et « **termes spécifiques** ». Chaque sous-commande correspond à une étape importante de notre premier module « **Analyse** ». L'utilisateur passe par ces étapes pour extraire respectivement tous les termes simples, les termes simples fréquents, les termes composés, les termes composés fréquents et les termes spécifiques de la visioconférence. Ces extractions seront réalisées en appliquant les algorithmes d'extractions des termes cités dans le chapitre modélisation. Le résultat sera affiché dans trois tables nommées respectivement « **termes simples** », « **termes composés** » et « **termes spécifiques** ».

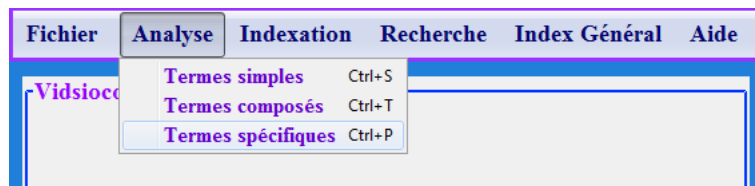


Figure 3. 5 : Extraction des termes

Pour extraire les termes spécifiques, nous proposons une interface sous forme des check boxes permettant aux utilisateurs de cocher les termes de poids faibles et riches sémantiquement. La **figure 3.6** illustre cette interface.

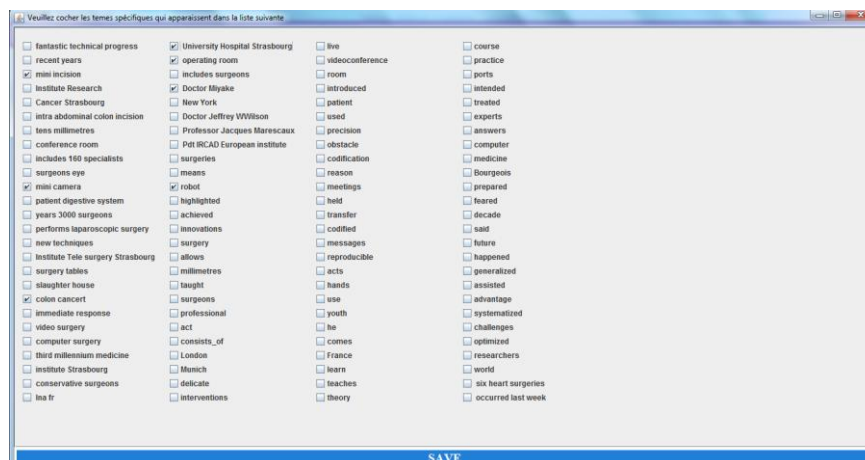


Figure 3. 6 : Extraction des termes spécifiques

Les listes correspondant respectivement aux termes seront affichées dans une table de trois colonnes nommées « **Liste des termes simples** », « **Liste des termes composés** » et « **Liste des termes spécifiques** ». Les colonnes « **Liste des termes simples** » et « **Liste des termes composés** » sont divisées respectivement en deux : la première partie affiche tous les termes trouvés (respectivement simples et composés) et la deuxième partie affiche les termes (respectivement simples et composés) fréquents.

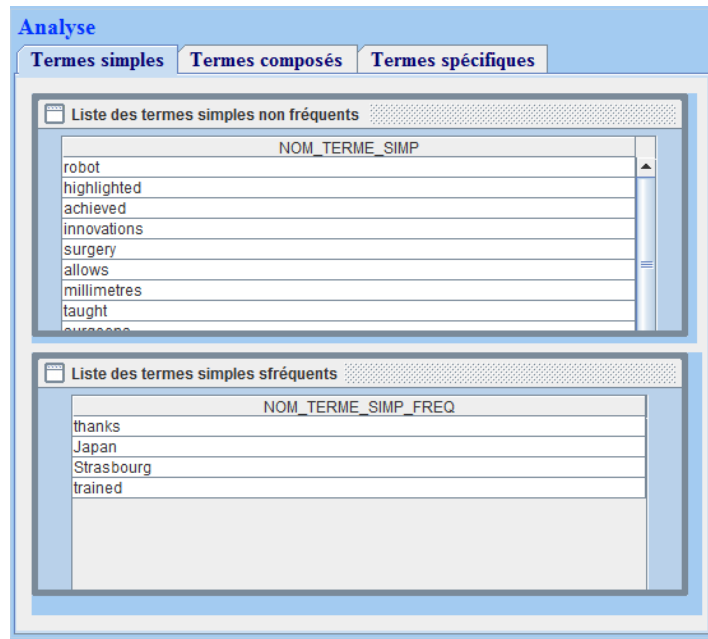


Figure 3. 7 : Liste des termes simples

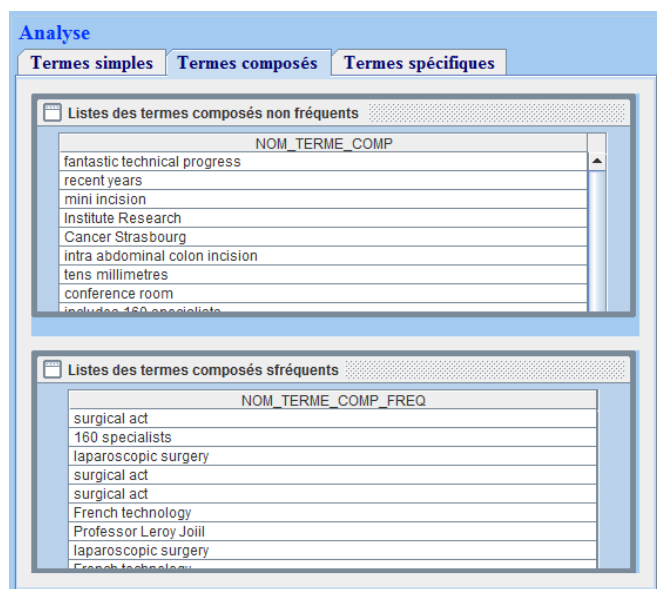


Figure 3. 8 : Liste des termes composés

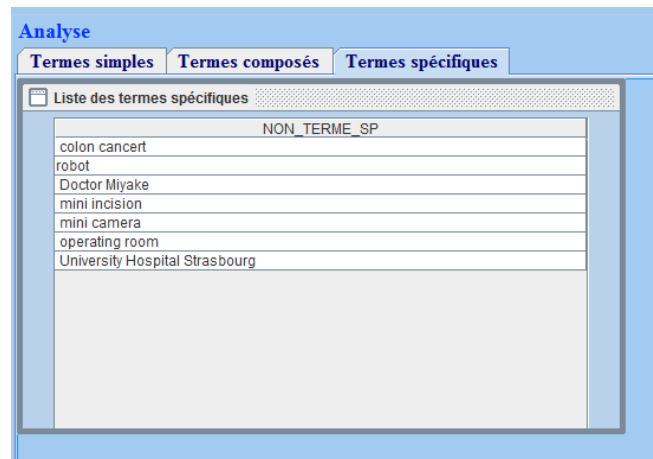


Figure 3. 9 : Liste des termes spécifiques

Les termes simples fréquents, les termes composés fréquents et les termes spécifiques seront ajouté automatiquement à la **liste des termes** (cf. **Figure 3.10**). Cette liste sera utilisée, par la suite pour trouver les concepts.

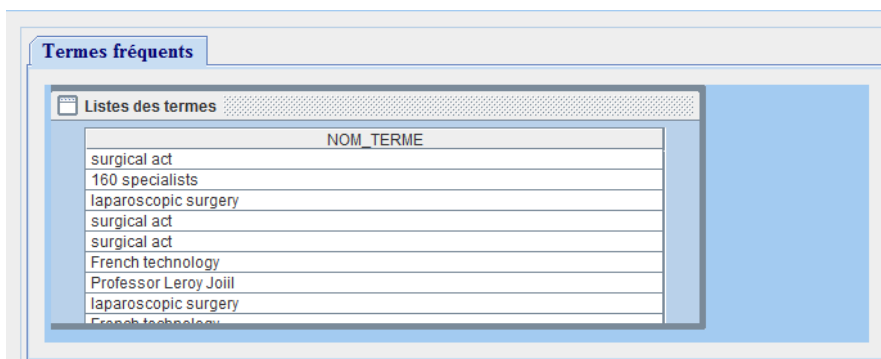


Figure 3. 10 : Liste des termes

II.2.2. Module d'Indexation

Nous avons déjà présenté l'architecture générale de notre module d'indexation dans le deuxième chapitre de la deuxième partie. L'outil proposé offre une interface graphique permettant de paramétrer le module suivi pour réaliser le module d'indexation proposé.

- 1- Extraction des concepts en identifiant les concepts associés aux termes simples ou termes composés après la projection de ces derniers sur les ressources sémantiques (cf. **Algorithme 3**).
- 2- Extraction des relations en cherchant les éventuelles relations sémantiques qui existent entre les concepts présents dans une même phrase. Ces relations doivent figurer dans l'une des ressources (cf. **Algorithme 4**).
- 3- Affichage de l'index sous forme d'un graphe conceptuel reliant les concepts dégagés dans l'étape précédente par les relations sémantiques.

Sous le menu « **Indexation** », nous trouvons les commandes « **Concepts** », « **Relation sémantiques** » et « **Index** ». La commande « **Concepts** » permet d’extraire les concepts relatifs aux termes fréquents de la visioconférence obtenus lors du module d’Analyse tout en utilisant une ou plusieurs ontologie(s). La commande « **Relations sémantiques** » permet de trouver les différentes relations sémantiques existantes entre les concepts.

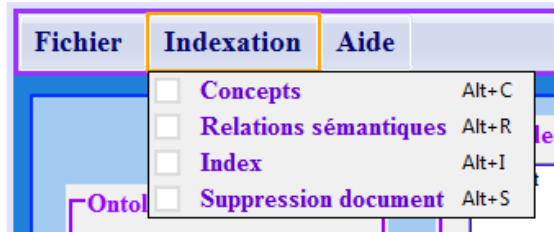


Figure 3. 11 : Menu « Indexation »

Le menu Fichier contient deux commandes : « **Importer** » et « **Quitter** ». La commande « **Importer** » permet d’importer une ontologie de domaine selon le choix de l’utilisateur.

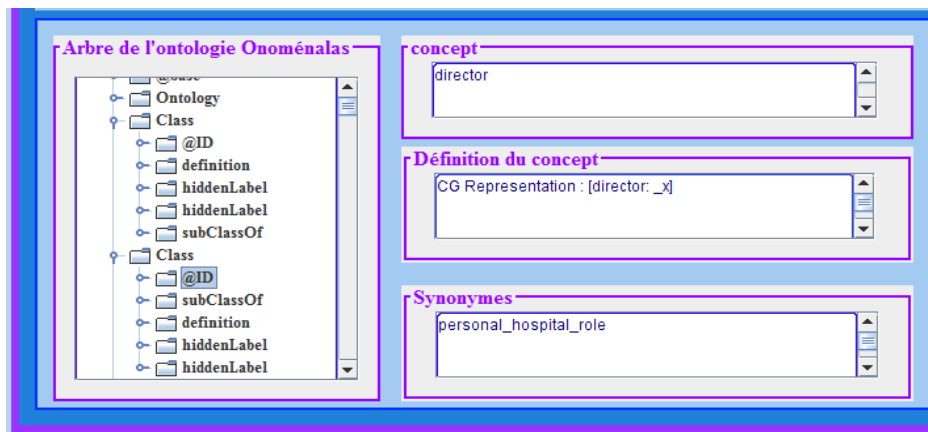


Figure 3. 12 : Ontologie affiché sous la forme arborescente

Les listes correspondantes respectivement aux concepts et aux relations conceptuelles qui existent entre ces concepts et figurent dans la liste des relations sémantiques de l’ontologie, seront affichées dans deux tables d’une colonne chacune nommées respectivement « **Liste des concepts** » et « **Liste des relations sémantiques** » en cliquant respectivement sur les sous commandes « **Concepts** » et « **Extraction des relations sémantiques** ». La table « **Liste des concepts** » est divisée en deux parties : la première affiche les concepts trouvés dans l’ontologie comme des concepts valides et la deuxième, affiche les autres concepts considérés comme des nouveaux concepts (cf. **Figure 3.13**). L’extraction des concepts et des relations sémantiques sera réalisée par les algorithmes d’extraction des concepts et des relations sémantiques cités dans le deuxième chapitre de la deuxième partie.

La commande « **Index** » du menu « **Indexation** » permet d’afficher les concepts et les relations sémantiques sous forme d’un graphe conceptuel. Les concepts sont présentés entre

deux accolades « [nom_concept] » et les relations entre deux parenthèses« (nom_relation) ». Les concepts et les relations sémantiques sont reliés entre eux par des flèches (→). La **figure 3.15** illustre le graphe conceptuel relatif à la visioconférence « Vidéo-chirurgie à Strasbourg ».

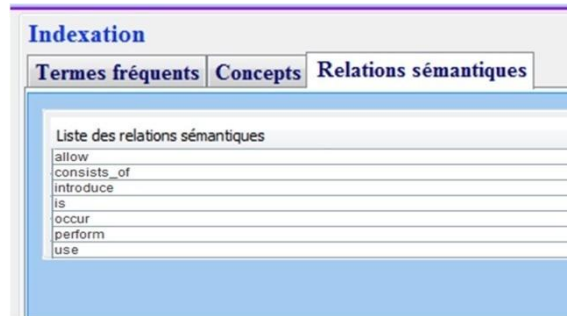
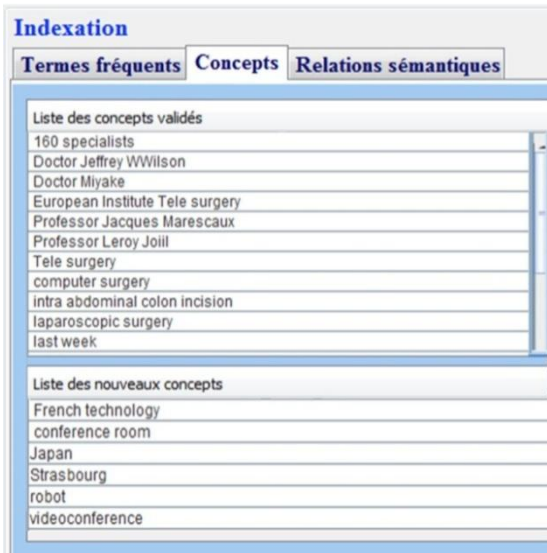


Figure 3. 13 : Liste des relations sémantique entre les concepts

Figure 3. 14 : Liste des concepts

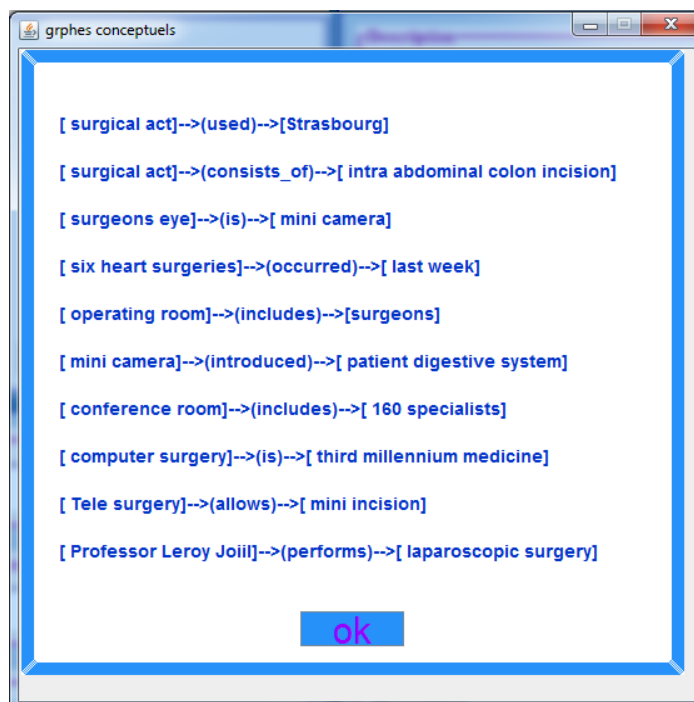


Figure 3. 15 : Index sous forme de graphe conceptuel.

La fusion des différents index relatifs aux visioconférences de notre corpus présente un index général du corpus.

La mise à jour de l'index général sera effectuée selon les algorithmes 5, 6, 7 et 8. Cette mise à jour n'est réalisée que sur la visioconférence concernée tout en conservant la cohérence entre le corpus et l'index pour trouver les documents les plus pertinents.

Par exemple, la suppression d'un document déjà indexé permet la suppression de son index de la base d'index. Cela se fait par la spécification de son chemin de sauvegarde, (cf. **Figure 3.16**).

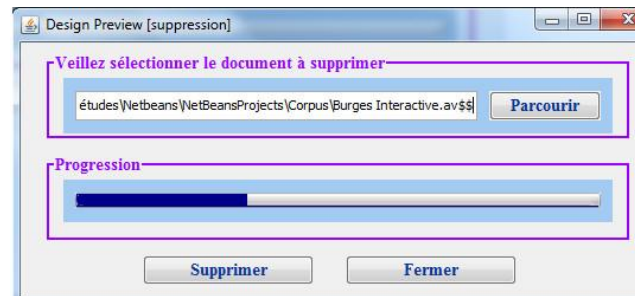


Figure 3. 16 : Fenêtre de suppression d'un document dans la base d'index.

II.2.4. Module de Recherche

Le module de recherche d'OSSVIRI est composé de deux étapes : reformulation de la requête et appariement document-requête. L'utilisateur écrit sa requête en utilisant ses propres termes. Cette requête sera, par la suite, traitée et analysée par l'étape de reformulation. Le système procède à l'extraction de l'ensemble des termes à rajouter afin de produire une nouvelle requête. Il récupère les concepts spécifiques et les relations sémantiques entre eux en utilisant les ontologies de domaine. Ces concepts et ces relations seront stockés dans une base relationnelle sous forme d'un index. Ce dernier sera envoyé, par la suite, à la deuxième étape de notre processus « Appariement ». L'étape de reformulation utilise aussi le profil utilisateur s'il existe. Ce profil sert à spécifier de plus le besoin de l'utilisateur.

Avant d'écrire sa requête, l'utilisateur a le choix de s'identifier ou non. Dans le premier cas, s'il a un compte, il suffit de se connecter avec son nom utilisateur et son mot de passe (cf. **Figure 3.18**). Dans le cas inverse (cf. **Figure 3.19**), il peut s'inscrire pour avoir un compte. Et dans le deuxième cas, l'utilisateur peut accéder au module de recherche comme un invité. Dans ce cas, l'étape de reformulation n'utilise pas le PU.

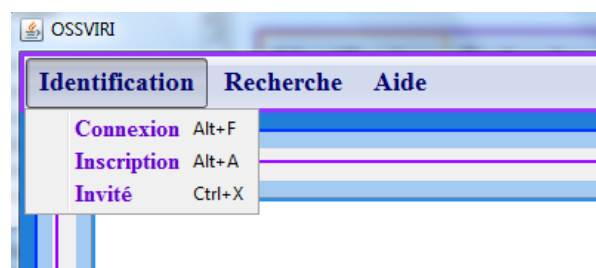


Figure 3. 17 : Identification de l'utilisateur

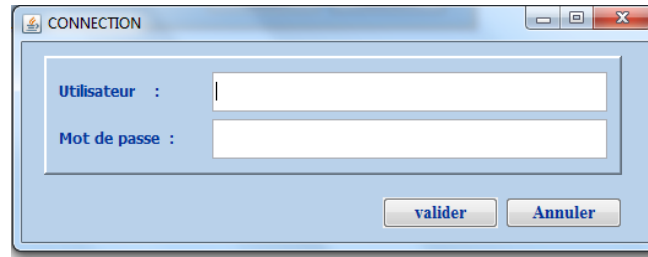


Figure 3. 18 : Connexion de l'utilisateur

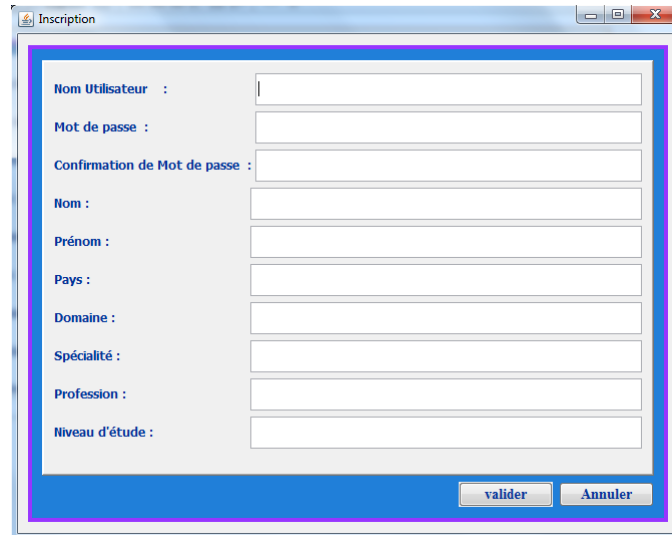


Figure 3. 19 : Inscription de l'utilisateur

Dans l'étape de l'appariement, le système compare l'index de la requête à ceux de la visioconférence du corpus en se basant sur l'algorithme 11. Les visioconférences ayant un score de similarité supérieur à un seuil fixé à l'avance seront ordonnées selon leur score de pertinence calculé par l'algorithme 12.

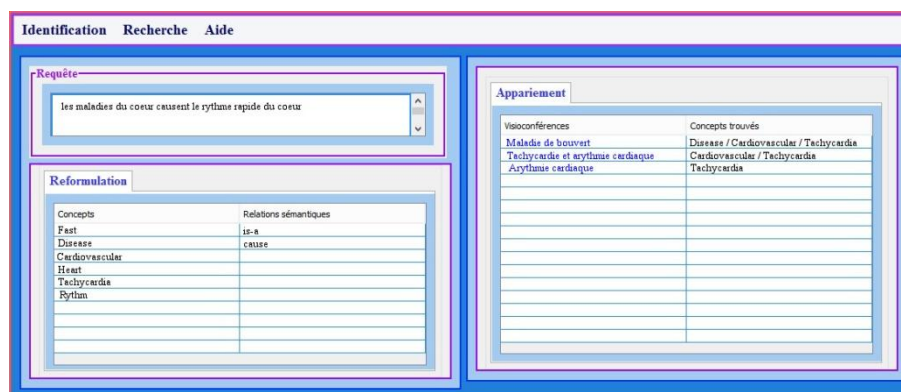


Figure 3. 20 : Menu recherche

Les résultats obtenus, après l'étape d'appariement, seront affichés à l'utilisateur sous forme d'une table à deux colonnes. La première colonne contient les noms des visioconférences et la deuxième présente les concepts spécifiques à chaque visioconférence (*cf.* **Figure 3.20**).

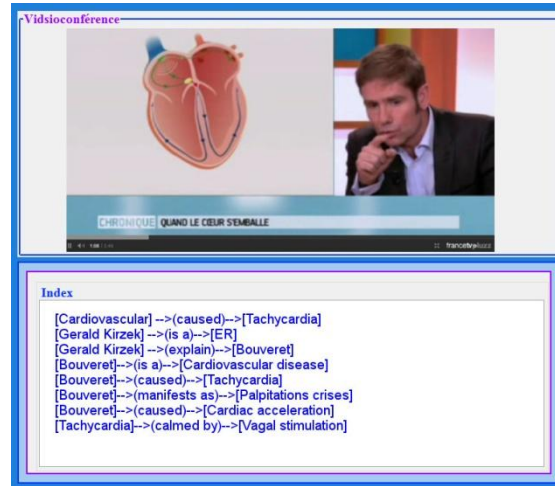


Figure 3. 21 : Résultat du module de recherche.

Conclusion

Dans ce chapitre, nous avons présenté l'implémentation de notre système de recherche d'information « OSSVIRI ». Nous avons tout d'abord présenté l'environnement de développement ainsi que les différents outils utilisés, puis nous avons donné une description détaillée du système « OSSVIRI » à travers des fenêtres de capture qui représentent les interfaces de ce dernier. Ces fenêtres sont conçues de manière à être conviviales et simples à utiliser. Cette étape nous a aussi permis de nous familiariser avec les outils utilisés pour le développement du système « OSSVIRI ».

Chapitre 2 : Création du corpus et évaluation

Ce chapitre présente la partie expérimentale de notre travail. Nous décrivons la série de tests utilisés dans nos expérimentations. Nous analysons, ensuite, les résultats obtenus à partir de notre système.

Dans le domaine de RI, les approches sont évaluées sur des collections de tests et en comparant leurs résultats à ceux obtenus par d'autres systèmes. Ces collection sont constituées d'un corpus de documents et d'un ensemble de requêtes pour le quels nous connaissons l'ensemble des documents pertinents dans le corpus de la collection.

Introduction	164
I.Descriptif du protocole d'expérimentation	164
II.Construction du corpus	165
II.1. Corpus du test.....	165
II.2. Corpus d'appui	165
II.3. Ressources externes	166
II.3.1. Présentation.....	166
II.3.2. Les concepts dans les ressources	166
II.3.3. Les relations entre les concepts et les types sémantiques	166
II.3.4. Les relations sémantiques entre les types sémantiques	167
III.Expérimentations et évaluations.....	167
III.1. Système OSSVIRI.....	168
III.2. Méthodologie d'évaluation.....	168
III.2.1. Description générale de la méthode d'évaluation.....	168
III.2.2. Mesure d'évaluation	168
III.3. Module « Analyse ».....	170
III.3.1. Extraction des termes simples par élimination des mots vides	170
III.3.2. Extraction des termes composés : détermination du seuil IMA	171
III.4. Module « Indexation »	172
III.4.1. Extraction des concepts.....	172
III.4.2. Extraction des relations sémantiques.....	173
III.4.3. Traitement des ambiguïtés	174
III.5. Module « Recherche ».....	175
III.5.1. Reformulation des requêtes	175
III.5.2. Appariement Document-Requête	176
Conclusion	181

Introduction

Nous décrivons, dans ce chapitre, l'expérimentation que nous avons menée au cours de cette thèse. Son objectif est l'évaluation de la qualité des résultats obtenus à travers notre système.

Afin de valider notre système, nous avons choisi une base de taille importante de documents visioconférences médicales. L'évaluation s'est réalisée en deux phases :

- la construction du corpus de test ;
- l'évaluation du résultat sur le corpus.

La démarche adoptée pour chaque étape et les choix associés sont détaillés dans la suite de ce chapitre.

I. Descriptif du protocole d'expérimentation

La **figure 3.22** illustre le protocole d'expérimentation.

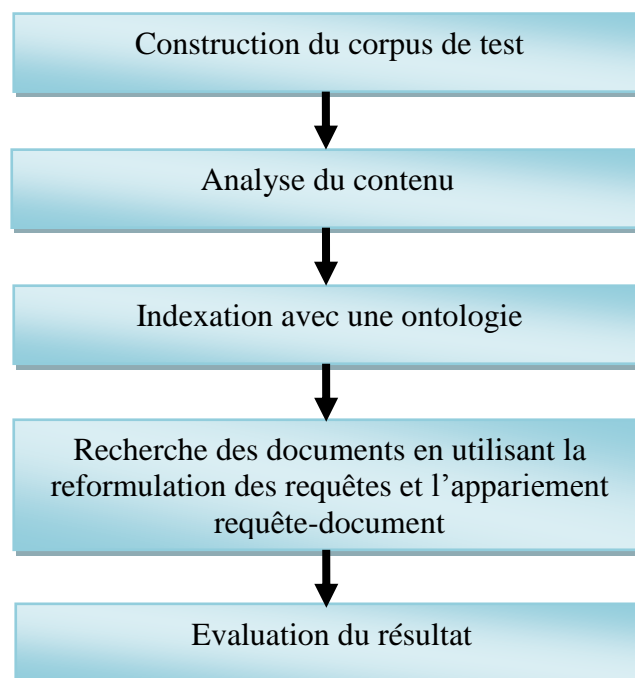


Figure 3.22 : Le protocole d'expérimentation

Les différentes étapes de notre expérimentation sont les suivantes :

1. construction d'un corpus qui répond à nos besoins, en l'occurrence des visioconférences médicales ;
2. module d'analyse du contenu des visioconférences pour les rendre accessibles et exploitables ;
3. module d'indexation des visioconférences analysées en utilisant des ontologies de domaine. Les résultats obtenus sont sauvegardés dans une base d'index sous forme des GCs ;

4. module de recherche qui consiste à calculer les mesures de correspondance entre les requêtes reformulées et les index des visioconférences pour bien satisfaire les besoins de l'utilisateur ;
5. évaluation du résultat.

II. Construction du corpus

II.1. Corpus du test

Nous expérimentons notre approche sur un corpus composé d'une centaine de visioconférences médicales archivées. À chaque visioconférence est associée une description textuelle écrite en langue anglaise. Cette description est présentée sous forme d'un fichier XML (cf. **Figure 3.23**).

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<?mso-application progid="Word.Document"?>
<w:wordDocument xmlns:aml="http://schemas.microsoft.com/aml/2001/core" xmlns:dt="uuid:C2F41010-65B3-11d1-A29F-00AA00C14882"
xmlns:ve="http://schemas.openxmlformats.org/markup-compatibility/2006" xmlns:o="urn:schemas-microsoft-com:office:office" xmlns:v=
"urn:schemas-microsoft-com:vml" xmlns:w10="urn:schemas-microsoft-com:office:word" xmlns:w=
"http://schemas.microsoft.com/office/word/2003/wordml" xmlns:wx="http://schemas.microsoft.com/office/word/2003/auxHint" xmlns:wsp=
"http://schemas.microsoft.com/office/word/2003/wordml/sp2" xmlns:sl="http://schemas.microsoft.com/schemaLibrary/2003/core"
xmlns:ns0="ActionsPane" xmlns:ns1="ActionsPane3" w:macrosPresent="no" w:embeddedObjPresent="no" w:ocxPresent="no" xml:space=
"preserve"><w:ignoreSubtree w:val="http://schemas.microsoft.com/office/word/2003/wordml/sp2"/><o:DocumentProperties><o:Author>
AMENI</o:Author><o:LastAuthor>AMENI</o:LastAuthor><o:Revision>4</o:Revision><o:TotalTime>390</o:TotalTime><o:Created>
2012-09-30T14:18:00Z</o:Created><o:LastSaved>2012-09-30T21:00:00Z</o:LastSaved><o:Pages>1</o:Pages><o:Words>127</o:Words>
<o:Characters>704</o:Characters><o:Lines>5</o:Lines><o:Paragraphs>1</o:Paragraphs><o:CharactersWithSpaces>830</
o:CharactersWithSpaces><o:Version>12</o:Version></o:DocumentProperties><w:fonts><w:defaultFonts w:ascii="Calibri" w:fareast=
"Calibri" w:h-ansi="Calibri" w:cs="Arial"/><w:font w:name="Times New Roman"><w:panose-1 w:val="02020603050405020304"/><w:charset
w:val="00"/><w:family w:val="Roman"/><w:pitch w:val="variable"/><w:sig w:usb-0="E0002AFF" w:usb-1="C0007841" w:usb-2="00000009"
w:usb-3="00000000" w:csb-0="000001FF" w:csb-1="00000000"/></w:font><w:font w:name="Arial"><w:panose-1 w:val="020B06040202020204"
/><w:charset w:val="00"/><w:family w:val="Swiss"/><w:pitch w:val="variable"/><w:sig w:usb-0="E0002AFF" w:usb-1="C0007843" w:usb-2=
"00000009" w:usb-3="00000000" w:csb-0="000001FF" w:csb-1="00000000"/></w:font><w:font w:name="Courier New"><w:panose-1 w:val=
"02070309020205020404"/><w:charset w:val="00"/><w:family w:val="Modern"/><w:pitch w:val="fixed"/><w:sig w:usb-0="E0002AFF" w:usb-1=
"C0007843" w:usb-2="00000009" w:usb-3="00000000" w:csb-0="000001FF" w:csb-1="00000000"/></w:font><w:font w:name="Symbol">
<w:panose-1 w:val="05050102010706020507"/><w:charset w:val="02"/><w:family w:val="Roman"/><w:pitch w:val="variable"/><w:sig
w:usb-0="00000000" w:usb-1="10000000" w:usb-2="00000000" w:usb-3="00000000" w:csb-0="80000000" w:csb-1="00000000"/></w:font>
<w:font w:name="Wingdings"><w:panose-1 w:val="05000000000000000000"/><w:charset w:val="02"/><w:family w:val="auto"/><w:pitch w:val=
"variable"/><w:sig w:usb-0="00000000" w:usb-1="10000000" w:usb-2="00000000" w:usb-3="00000000" w:csb-0="80000000" w:csb-1=
"00000000"/></w:font><w:font w:name="Cambria Math"><w:panose-1 w:val="02040503050406030204"/><w:charset w:val="01"/><w:family
w:val="Roman"/><w:notTrueType/><w:pitch w:val="variable"/><w:sig w:usb-0="00000000" w:usb-1="00000000" w:usb-2="00000000" w:usb-3=
"00000000" w:csb-0="00000000" w:csb-1="00000000"/></w:font><w:font w:name="Calibri"><w:panose-1 w:val="020F0502020204030204"/>
```

Figure 3. 23 : Exemple d'un document XML

II.2. Corpus d'appui

Afin d'extraire les mots vides, le module d'analyse se base sur la distance intertextuelle du domaine. Il s'agit de déterminer le lexique commun de deux corpus de domaine disjoints. Le premier corpus est le corpus du test et le deuxième est le corpus d'appui. Dans nos expérimentations, nous utilisons le corpus du parlement européen⁶ anglais-français comme corpus d'appui.

⁶<http://www.statmt.org/europar1/>

II.3. Ressources externes

Dans cette section, nous présentons les ressources sémantiques OntoMénélas et UMLS qui ont été choisies pour nos expérimentations. Ces ressources nous permettent dans un premier temps d'identifier les concepts associés aux termes et dans un deuxième temps d'identifier les relations entre ces concepts.

II.3.1. Présentation

Les deux ressources choisies couvrent une grande partie du domaine médical. Elles ont été créées dont le but de faciliter la recherche et l'intégration d'informations provenant des sources d'information médicales électroniques. Elles définissent l'organisation des concepts et des relations entre eux. Ces concepts sont organisés en classes. À chaque concept est associée au moins une classe. Ces classes forment des types sémantiques. Ces types sont reliés entre eux par des relations sémantiques. Dans la suite, nous décrivons les concepts, les types sémantiques et les relations entre ces types que nous utilisons dans nos expérimentations.

II.3.2. Les concepts dans les ressources

Dans les ressources, chaque concept est identifié par un identificateur unique appelé Identificateur Unique du Concept (IUC). Il est relié à un ensemble de termes. Ces derniers dénotent le concept et ils sont présentés par des Terms Associés au Concept (TAC). Les TAC sont liés à un Indicateur Préféré ou Non (IPN) qui indique si le terme est préféré (Y) ou non (N). Pour chaque concept, sa source est mentionnée par une Ontologie Source (OS) cf. **tableau 3.1**.

Tableau 3. 1: Exemple de concept C0001365

IUC	IPN	TAC	OS
C0001365	Y	Cerebrovascular Disease	UMLS
C0001365	N	Heart Rhythm	OntoMénélas
C0001365	N	CVa	UMLS

II.3.3. Les relations entre les concepts et les types sémantiques

Dans l'ontologie, les concepts sont organisés en types sémantiques. Dans l'UMLS, nous avons trouvé 135 types. Chaque type est identifié par un identificateur unique appelé Identificateur Unique du Type Sémantique (IUTS). Les différents types sémantiques reliés au concept C1516400 sont présentés dans le **tableau 3.2**.

II.3.4. Les relations sémantiques entre les types sémantiques

Les types sémantiques sont reliés entre eux par des relations sémantiques. Ces relations seront utilisées, dans nos travaux, pour extraire les relations sémantiques entre les concepts.

Pour deux concepts C1 et C2, nous pouvons trouver une relation sémantique si seulement si nous trouvons une relation sémantique (RS) entre le type associé à C1 et celui associé à C2. Ce type de relation est présenté par le **tableau 3.3**.

Tableau 3. 2: Le concept C1516400 et ses types sémantiques

IUC	IUTS
C1516400 (Biomedical Occupation or Discipline)	T83 (Organism Attribute : Clinical Attribute)
C1516400 (Biomedical Occupation or Discipline)	T85 (Intellectual Product : Classification)
C1516400 (Biomedical Occupation or Discipline)	T86 (Intellectual Product: Regulation or Law)
C1516400 (Biomedical Occupation or Discipline)	T87 (Language)
C1516400 (Biomedical Occupation or Discipline)	T91 (Organization : Health Care Related Organization)
C1516400 (Biomedical Occupation or Discipline)	T92 (Organization : Professional Society)
C1516400 (Biomedical Occupation or Discipline)	T93 (Organization : Self-help or Relief Organization)
C1516400 (Biomedical Occupation or Discipline)	T96 (Group : Professional or Occupational)
C1516400 (Biomedical Occupation or Discipline)	T100 (Group : Patient or Disabled Group)

Tableau 3. 3: Une relation sémantique entre deux types sémantiques

IUTS	RS	IUTS
T27 (Enzyme)	T (is-a)	T42 (Biologically Active substance)

III. Expérimentations et évaluations

Dans cette section, nous présentons les expérimentations effectuées pour évaluer notre système. En premier lieu, nous testons le module d'analyse. Cette partie contient les prétraitements des corpus et des requêtes tout en extrayant les termes simples et les termes composés qui les constituent. En deuxième lieu, nous évaluons le module d'indexation. Nous vérifions les résultats d'extraction des concepts et des relations sémantiques. Et en troisième lieu, nous évaluons notre module de recherche avec ces deux parties : reformulation des requêtes et appariement document-requête. Mais avant de traiter ces trois parties, nous

présentons l'architecture générale du système développé. Nous terminons par la présentation des métriques utilisées, la représentation intermédiaire ainsi que la fonction de correspondance et l'illustration des résultats de nos expérimentations.

III.1. Système OSSVIRI

Afin de valider notre système et montrer sa faisabilité, nous avons réalisé un ensemble d'expérimentations. OSSVIRI constitue un SRI qui permet d'indexer les visioconférences médicales et de rechercher les informations utiles. Le système proposé offre une interface graphique permettant de paramétrer les différents modules de notre système comme c'est déjà mentionné dans le chapitre précédent. L'architecture d'OSSVIRI est présentée dans la **figure 3.24**.

III.2. Méthodologie d'évaluation

Dans cette partie, nous décrivons, en premier lieu, la méthode d'évaluation utilisée. En deuxième lieu, nous exposons les mesures trouvées. Et, enfin, nous présentons le SRI de référence.

III.2.1. Description générale de la méthode d'évaluation

Afin d'étudier la performance de notre SRI, nous comparons l'efficacité de notre système à celle des systèmes existants basée sur les analyses statistiques.

III.2.2. Mesure d'évaluation

Nous utilisons les mesures de performance des SRI pour évaluer notre système.

III.2.2.1. Les mesures Rappel / Précision

Ces deux mesures sont obtenues en partitionnant l'ensemble des objets (documents / concepts/mots vides) restitués par le SRI en deux catégories « pertinent/utile » et « non pertinent/non utile ». Elles sont définies comme suit :

$$\mathbf{Rappel} = \frac{a}{a+c} \quad \text{et} \quad \mathbf{Précision} = \frac{a}{a+b} \quad (3.1)$$

Où

- a est le nombre d'objets utiles retrouvés
- b est le nombre d'objets non utiles retrouvés
- c est le nombre d'objets utiles non retrouvés

III.2.2.2. La courbe Rappel / Précision

Cette courbe est souvent utilisée comme mesure combinée d'évaluation des SRI. Elle donne la valeur de précision correspondante à chaque point de rappel.

III.2.2.3. Précision à x documents

Il s'agit de la valeur de la précision obtenue pour les x premiers objets de la liste ordonnée des objets restitués. Il est noté P@x.

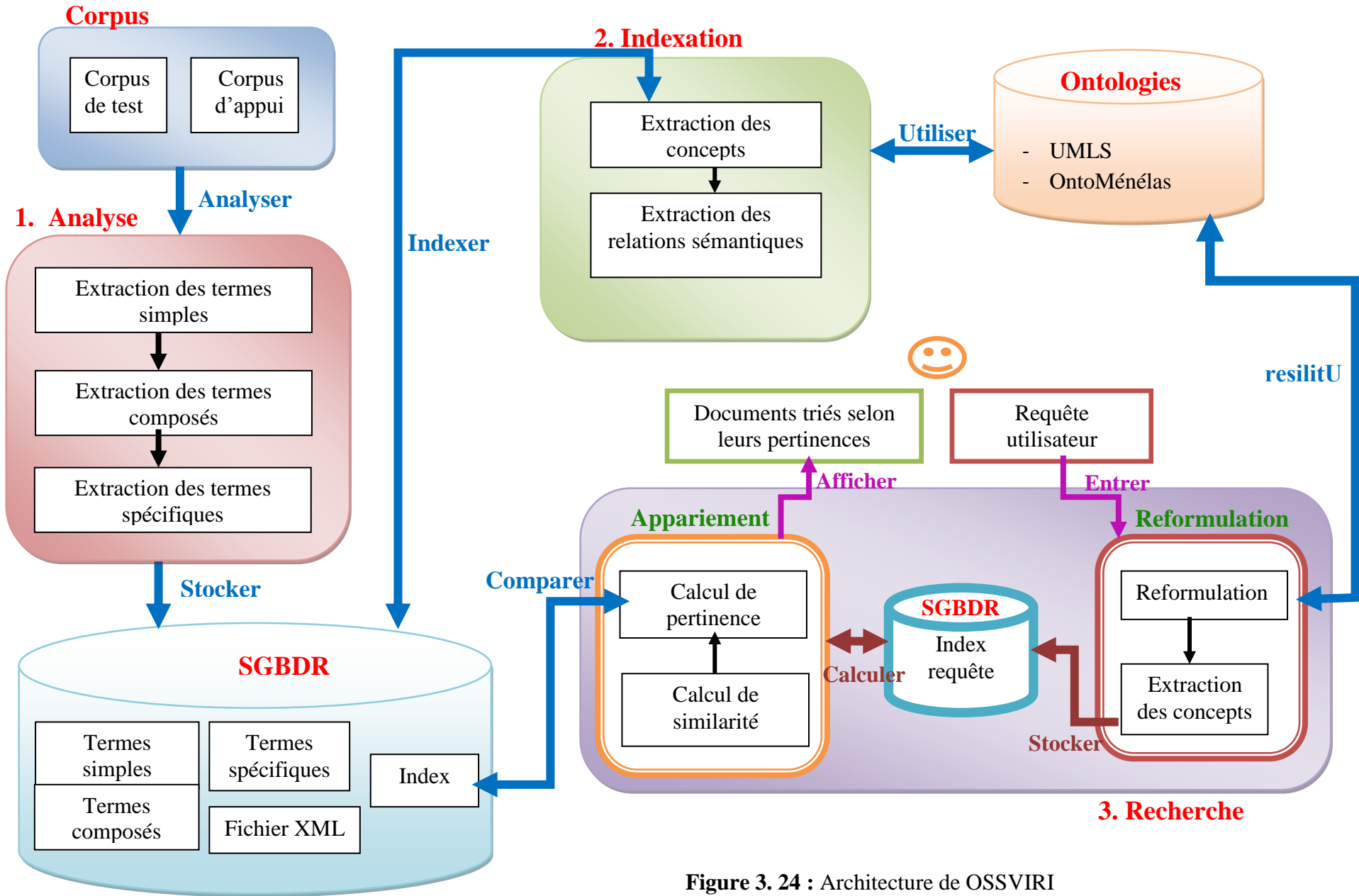


Figure 3. 24 : Architecture de OSSVIRI

III.2.2.4. Précision Moyenne Globale

La précision moyenne est la moyenne des valeurs de précision calculée à chaque objet utile de la liste ordonnée. Si un objet utile est retourné à la $x^{\text{ième}}$ position, la précision pour cet objet est « la précision à x documents ». Si un objet utile n'a pas été trouvé par le système, la précision pour cet objet est nulle. La MAP se calcule comme suit :

$$MAP = \sum_{i=1}^{N_q} \frac{P_i(X)}{N_q} \quad (3.2)$$

Où :

- N_q est le nombre total d'objets
- $P_i(X)$ donne la précision du $i^{\text{ème}}$ objet correspondante au taux de rappel X .

III.3. Module « Analyse »

III.3.1. Extraction des termes simples par élimination des mots vides

Afin d'éliminer les mots vides, nous projetons le corpus d'appui sur les fichiers XML de notre collection. La liste des mots obtenue forme la liste des termes simples. Nous évaluons l'extraction des termes simples à travers l'évaluation de l'élimination des mots vides. Pour réaliser cette évaluation, nous utilisons la liste de références des mots vides de la langue anglaise. Cette liste est disponible sur le web⁷. Elle contient 116 mots en français et 179 mots en anglais.

La mesure rappel et la mesure précision utilisées dans notre évaluation sont définies comme suit :

$$\text{Précision} = \frac{\text{nombre de mots vides extraits et présents dans la liste de référence}}{\text{nombre de mots vides extraits}}$$

$$\text{Rappel} = \frac{\text{nombre de mots vides extraits et présents dans la liste de référence}}{\text{nombre de mots vides de la liste de référence}}$$

III.3.1.1. Résultats

$$\text{Précision} = \frac{336}{336} = 100\% \quad \text{Rappel} = \frac{336}{353} = 95.18\%$$

III.3.1.2. Synthèse

Dans nos expériences, nous avons extrait 336 mots vides. Ces mots sont présents dans la liste de références. La valeur de précision est égale à 100%. Mais certains mots vides ne sont pas extraits par le processus d'extraction. D'où la valeur de rappel est 95.18%. Cette valeur peut être expliquée par le fait que ces mots sont absents dans notre corpus comme « therefore » et « seeming ». Nous signalons que dans UMLS, certains mots vides comme « of », « in » et « the » sont associés à des concepts.

Cette évaluation montre que notre algorithme d'extraction des mots simples est efficace.

⁷<http://www.ranks.nl/stopwords>

III.3.2. Extraction des termes composés : détermination du seuil IMA

Notre processus d'extraction des termes composés procède à l'extraction de nouveaux termes à partir de la liste des termes simples en utilisant une mesure statistique IMA. Pour chaque couple de terme, nous calculons la valeur de l'IMA. Si cette valeur est inférieure à une valeur seuil, les termes seront considérés comme un terme composé.

Pour déterminer la valeur seuil, nous évaluons, la précision moyenne et la précision à 5 documents à différentes valeurs de l'IMA. Après un nombre d'itérations, nous obtenons des mesures constantes. La première valeur de l'IMA qui provoque cette stabilité, sera considérée comme la valeur seuil.

Les valeurs MAP et P@5 calculées à différentes valeurs du seuil IMA sont présentées par le **tableau 3.4**, **figure 3.25** et **figure 3.26**.

Tableau 3. 4: Valeur de la MAP et de la P@5 en fonction du seuil de l'IMA

IMA	MAP	P@5	IMA	MAP	P@5
7	0.1018	0.0705	15	0.2421	0.0875
8	0.1541	0.0812	16	0.2421	0.0875
9	0.2572	0.0938	17	0.2421	0.0875
10	0.2499	0.0911	18	0.2421	0.0875
11	0.2454	0.0896	19	0.2421	0.0875
12	0.2435	0.0875	20	0.2421	0.0875
13	0.2407	0.0875	21	0.2421	0.0875
14	0.2421	0.0875			

Dans cette expérience, nous avons trouvé une amélioration du MAP et P@5 pour les valeurs du seuil IMA entre 7 et 14. Nous remarquons que les deux valeurs deviennent constantes à partir de la valeur 14. Cette valeur sera utilisée comme valeur seuil de l'IMA.

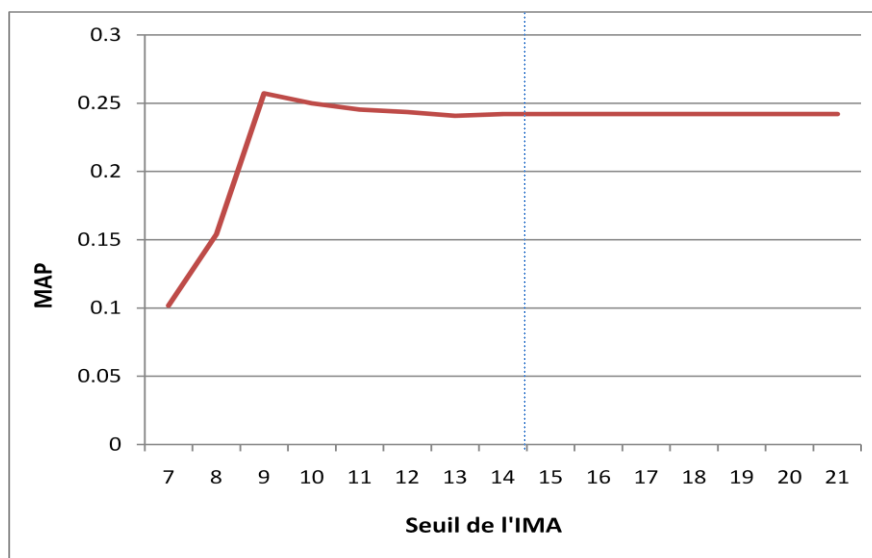


Figure 3. 25 : Variation de MAP en fonction du seuil de l'IMA

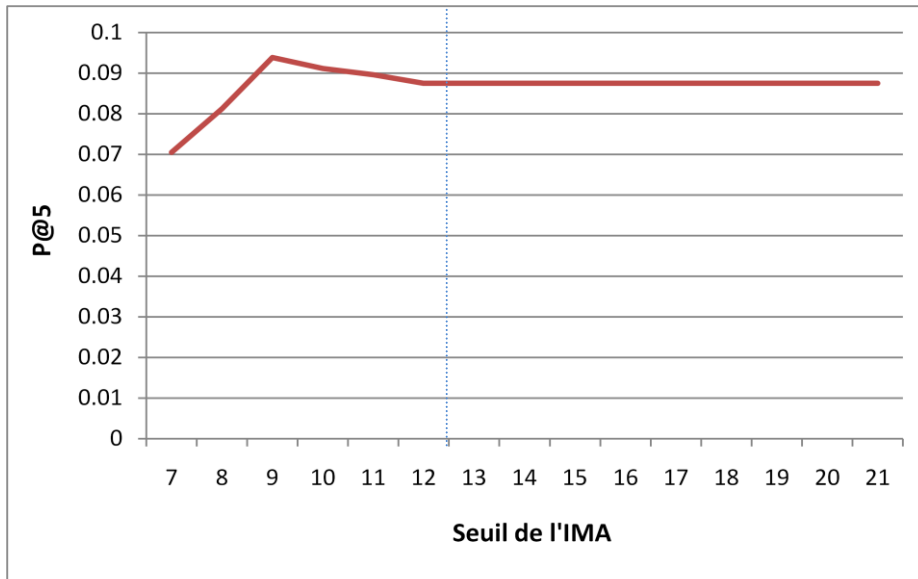


Figure 3. 26 : Variation de P@5 en fonction du seuil de l'IMA

III.4. Module « Indexation »

Pour évaluer notre module d'indexation, nous utilisons MAP pour avoir une vue globale de la performance de notre système et P@5 pour avoir un aperçu sur son efficacité sur les 5 premiers documents retournés par OSSVIRI.

III.4.1. Extraction des concepts

L'identification des concepts associés aux termes consiste à extraire les concepts à partir des documents de la collection en utilisant l'opérateur *Sc*. Ce dernier permet de déterminer l'ensemble des concepts dénotés par un terme.

III.4.1.1. Résultats

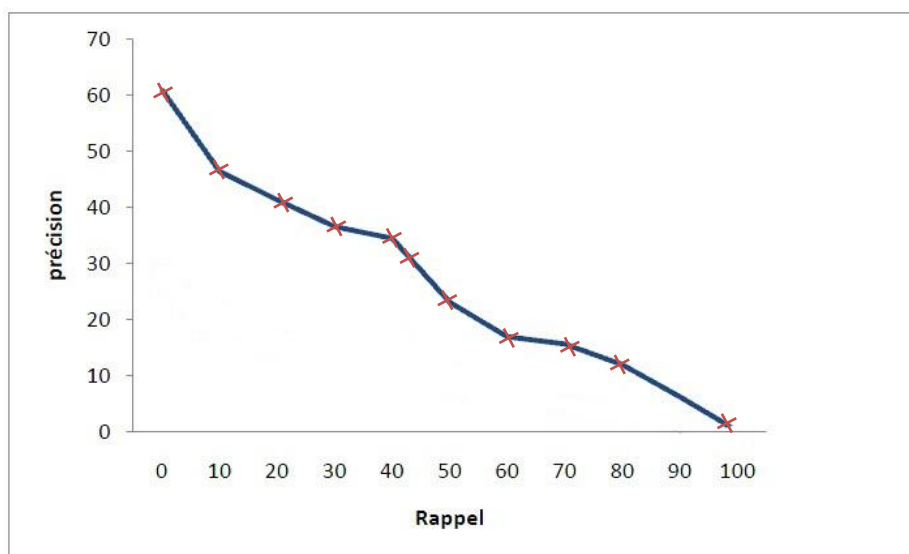


Figure 3. 27 : Courbe de la précision à 11 points de rappel

La courbe trouvée dans la **figure 3.27** correspond à la précision à 11 points de rappel de 50visioconférences.

III.4.1.2. Comparaison de nos résultats avec les résultats des approches existantes

Dans cette section, nous comparons les résultats de notre approche aux résultats obtenus par l'approche statistique la plus récente STAT de [Harrathi *et al.*, 11]. Cette comparaison est réalisée par la comparaison des valeurs des métriques MAP et P@5 obtenues par notre système OSSVIRI à ceux présentées dans STAT.

Tableau 3. 5: Résultat en MAP et de la P@5 pour les deux systèmes

Système	MAP	P@5	Δ MAP	Δ P@5
OSSVIRI	0.227	0.498		
STAT	0.244	0.425	- 6.967%	17.176%

Avec :

$$\Delta\text{MAP} = \frac{\text{MAP}(\text{OSSVIRI}) - \text{MAP}(\text{STAT})}{\text{MAP}(\text{STAT})} * 100 = \frac{0.227 - 0.244}{0.244} * 100 = -6.967$$

$$\Delta\text{P@5} = \frac{\text{P@5}(\text{OSSVIRI}) - \text{P@5}(\text{STAT})}{\text{P@5}(\text{STAT})} * 100 = \frac{0.498 - 0.425}{0.425} * 100 = 17.176$$

Cette comparaison montre que notre système donne des résultats meilleurs que l'approche STAT.

III.4.2. Extraction des relations sémantiques

L'OntoMénélas et l'UMLS sont composées par plusieurs classes dont chacune est associée à un type sémantique. Ces classes contiennent un ensemble de concepts liés entre eux par des relations sémantiques. Chaque concept peut appartenir à une ou plusieurs classe(s).

Dans nos travaux de thèse, nous avons utilisé l'hypothèse que deux concepts appartenant à une même phrase sont liés entre eux par une relation sémantique. Cette relation est détectée si :

1. un type sémantique TS1 associé à C1 ;
2. un type sémantique TS2 associé à C2 ;
3. TS1 et TS2 sont reliés par une relation sémantique.

Les concepts et les relations trouvés seront présentés sous forme de GC.

Pour détecter l'utilité des relations sémantiques dans l'indexation, nous comparons l'indexation par concepts seulement via l'indexation par concepts et relations sémantiques. Cette comparaison est illustrée dans la **figure 3.28**. Dans laquelle, la première courbe correspond à l'utilisation des concepts seuls pour indexer les documents du corpus. La deuxième courbe correspond à l'utilisation des concepts et des relations sémantiques pour indexer les documents. Elle donne des résultats meilleurs que ceux obtenus par la première

courbe. D'où, l'utilisation des relations en plus des concepts améliore la MAP et la P@5. Cette amélioration est détaillée dans le **tableau 3.6**.

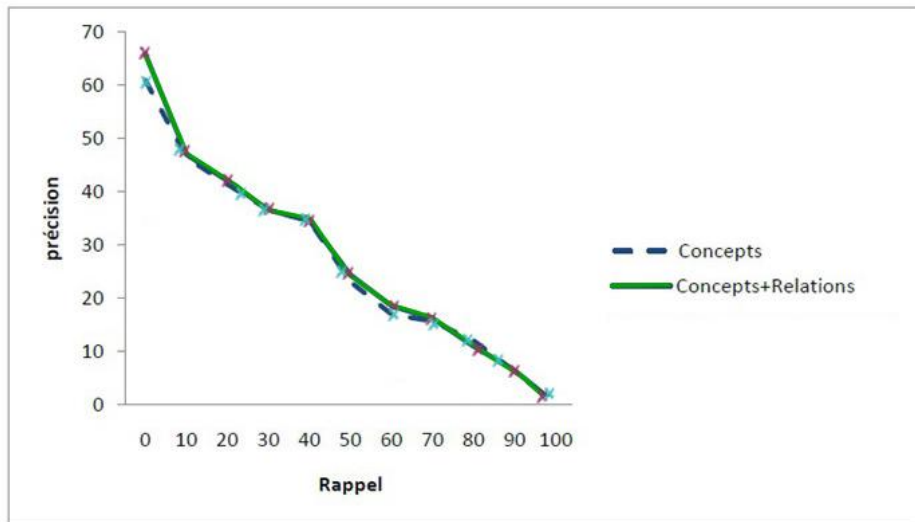


Figure 3. 28 : Courbe de la précision à 11 points de rappel : concepts / concepts + relations

Tableau 3. 6: Résultat en MAP et de la P@5 pour les deux scénarios : concepts / concepts + relations

Concepts seuls		Concepts + Relations		Δ MAP	Δ P@5
MAP	P@5	MAP	P@5		
0.254	0.465	0.261	0.489	2.756%	5.161%

III.4.3. Traitement des ambiguïtés

Durant cette étape, nous commençons par l'extraction des termes à partir des documents. Ces termes sont identifiés en utilisant l'opérateur *Sc*. Nous traitons l'ambiguïté sémantique. Nous favorisons les concepts qui possèdent des relations sémantiques avec d'autres concepts.

III.4.3.1. Résultats

Pour déterminer l'importance de la désambiguïsation des termes, nous comparons les résultats obtenus avec ou sans le traitement des ambiguïtés.

Les résultats obtenus sont présentés dans le **tableau 3.7** et le **tableau 3.8** et sont illustrés dans la **figure 3.29**.

Tableau 3. 7: Résultat en MAP et de la P@5 sans et avec traitements des ambiguïtés

Sans traitement d'ambigüité		Avec traitement d'ambigüité		Δ MAP	Δ P@5
MAP	P@5	MAP	P@5		
0.248	0.449	0.252	0.435	1.613%	-3.118%

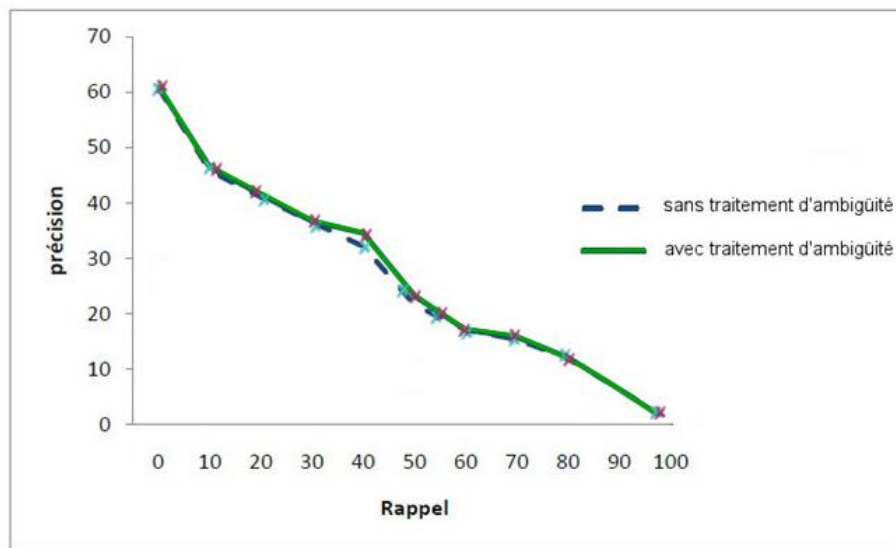


Figure 3. 29 : Courbe de la précision à 11 points de rappel : sans et avec traitements des ambiguïtés

Tableau 3. 8: Quelques résultats de la désambigüisation

	Sans traitement d'ambigüité	Avec traitement d'ambigüité	Δ NT
Nombre de termes non ambigus	3678991	4210338	14.443%
Nombre de termes associés à deux concepts	168833	174348	3.267%

III.4.3.2. Synthèse

En comparant les résultats d'extraction obtenus sans traitement d'ambigüité à ceux avec traitement d'ambigüité, nous avons trouvé une amélioration de la valeur MAP et une diminution de la valeur P@5. Les valeurs trouvées montrent que notre méthode de désambigüisation est performante et donne de bons résultats. Cette méthode a permis d'augmenter le nombre de termes non ambigus de 14.443% et de diminuer le taux d'ambigüité pour les autres termes à 3.267%.

III.5. Module « Recherche »

III.5.1. Reformulation des requêtes

Afin de mettre à l'épreuve les performances de module de recherche de notre système, nous avons mené une expérimentation qui vise à évaluer l'apport de l'utilisation de la sémantique via les ressources externes dans le processus de la reformulation des requêtes.

L'évaluation est faite en se basant sur la précision et le rappel. Ces derniers sont calculés de la manière suivante :

$$Rappel = \frac{\text{Nombre documents Pertinents retrouvés}}{\text{Nombre total documents pertinents}} \quad (3.3)$$

$$\text{Précision} = \frac{\text{Nombre documents Pertinents retrouvés}}{\text{Nombre total documents retrouvés}} \quad (3.4)$$

Notre module de recherche prend en entrée un ensemble de 50 requêtes en langage naturel. Pour juger l'importance de la reformulation des requêtes, nous examinons la pertinence des documents retournés à l'utilisateur avec la reformulation des requêtes en utilisant les ressources sémantiques et le profil utilisateur (R combinaison), avec :

- la reformulation en utilisant les ressources sémantiques uniquement (R ressource sémantique) ;
- la reformulation en utilisant le profil utilisateur uniquement (R profil utilisateur) ;
- et sans reformulation.

III.5.1.1. Résultats

Les résultats obtenus sont présentés dans le tableau suivant :

Tableau 3. 9: Résultat en Rappel et Précision sans et avec reformulation des requêtes

Type de recherche	Rappel (%)	Ecart Rappel	Précision (%)	Ecart Précision
R combinaison	92.20		90.37	
R ressource sémantique	85.10	7.10	80.89	9.48
R profil utilisateur	83.59	8.61	81.36	9.01
Sans reformulation	79.12	13.08	75.24	15.13

III.5.1.2. Synthèse

Les scores obtenus montrent que le taux de rappel est plus élevé dans le cas d'une recherche guidée par la RE et le PU. L'apport de la reformulation a produit un gain de 13,08 par rapport à la recherche sans reformulation des requêtes. Cet écart en termes de rappel s'explique par l'ajout de nouveaux termes à la requête. L'amélioration de 15,13 en termes de précision montre également que la recherche sémantique guidée par la RE et le PU est plus précise que la recherche guidée par la RE ou le PU et sans reformulation des requêtes. Ce taux signifie que peu de documents inutiles sont retournés par le système.

III.5.2. Appariement Document-Requête

III.5.2.1. Calcul de similarité

Pour étudier la qualité de l'ordonnancement de la liste des documents retournés dans l'ordre décroissant des scores de similarités avec la requête, nous examinons l'ensemble des n premiers documents retournés en utilisant la précision P_n ou le rappel R_n . Ces deux mesures sont notées respectivement $P@n$ et $R@n$. Les deux métriques ne sont pas indépendantes : quand l'une augmente, l'autre diminue.

La valeur de ces taux est influencée par le module d'indexation. En effet, plus l'indexation est exhaustive, plus le taux de rappel est potentiellement important : toutes les informations

susceptibles d'être pertinentes peuvent être restituées, mais certaines ne seront pas pertinentes pour l'utilisateur. Symétriquement, plus l'indexation est spécialisée, plus le taux de précision est élevé, mais cela induit le risque d'avoir un taux de rappel faible : les informations restituées seront pertinentes, mais d'autres informations pertinentes ne seront pas restituées.

La précision à 5, à 10, à 30 ... documents restitués présente des limites : par exemple si une requête donnée a seulement 6 documents pertinents et le système restitue ces 6 documents, la précision sera 0,6. Cette mesure n'illustre pas que tous les documents sont pertinents.

Pour pallier cette limite, il est important d'utiliser la R-précision. Cette dernière est la précision à n quand n est égal au nombre total de documents pertinents.

Pour évaluer notre processus d'appariement, nous avons choisi de comparer notre mesure de similarité *Sim* défini dans le chapitre 3 partie 2 à une autre mesure : la mesure *Cosinus*. Le nombre d'itérations effectuées est de 10.

La mesure Cosinus est définie par :

$$\mathbf{Cosinus} = \frac{\sum_i w_{di} * w_{qi}}{\sqrt{\sum_i w_{di}^2 * \sum_i w_{qi}^2}} \quad (3.5)$$

Nous comparons les deux mesures à l'aide des moyennes de précision à 5, à 10 et à 30 documents restitués, des moyennes de R-précision, des mesures F à 5, à 10 et à 30 documents restitués et la mesure MAP.

La mesure MAP est calculée par la formule suivante :

$$\mathbf{MAP} = \frac{1}{n} \sum_{i=1}^N \mathbf{P}(i) * \mathbf{R}(i) \quad (3.6)$$

avec :

- $\mathbf{R}(i) = 1$ si le $i^{\text{ème}}$ document restitué est pertinent ;
- $\mathbf{R}(i) = 0$ si le $i^{\text{ème}}$ document restitué est non pertinent ;
- $\mathbf{P}(i)$: la précision à i documents restitués ;
- n : le nombre de documents pertinents restitués ;
- N : le nombre total de documents.

La mesure F est une valeur synthétique qui exploite la précision et le rappel. Elle est calculée comme suit :

$$\mathbf{F} = (2 * \mathbf{Rappel} * \mathbf{Précision}) / (\mathbf{Rappel} + \mathbf{Précision}) \quad (3.7)$$

Pour évaluer l'ordonnement, nous calculons la mesure F_n à chaque rang n :

$$\mathbf{F@n} = (2 * \mathbf{R@n} * \mathbf{P@n}) / (\mathbf{R@n} + \mathbf{P@n}) \quad (3.8)$$

Nous avons utilisé cette mesure afin de trouver le rang n de la meilleure mesure F.

a. Résultats

Les mesures trouvées sont présentées par le **tableau 3.10**.

Les précisions ainsi que les mesures F vérifiées sont du même ordre de grandeur que celle calculée pour le *Cosinus*, tout en restant légèrement inférieure. Les conclusions immédiates à tirer sont que le *Sim* ne fournit pas une mesure suffisamment efficace dans le cadre de la RI : elle est légèrement inférieure à la mesure *Cosinus*, pour un temps de calcul très supérieur (le temps de calcul du *Cosinus* correspond au temps d'une itération du *Sim*, or il a fallu 10 itérations pour que l'algorithme converge vers les valeurs finales soit un coût temporel 10 fois supérieur) et une place mémoire également supérieure.

Tableau 3. 10: Précision à 5, à 10, à 30 documents restitués, R-précision moyenne et F à 5, à 10, à 30 documents pour *Sim* et *Cosinus*

Mesures	Cosinus	Sim	Δ
P@5	0,347	0,344	-0.8%
P@10	0,313	0,3	-4%
P@30	0,231	0,217	-6%
R-Prec	0,132	0,226	+40%
F@5	0,116	0,109	-6%
F@10	0,149	0,145	-2%
F@30	0,195	0,182	-7%

Avec

$$\Delta = \frac{\text{Sim} - \text{Cosinus}}{\text{Cosinus}} * 100$$

Pour avoir une idée générale en termes de précision et de rappel des résultats obtenus par les deux mesures, nous présentons la courbe précision-rappel en 11 points.

La **figure 3.30** montre que les deux mesures donnent des résultats similaires. Le *Cosinus* a les meilleurs scores pour un rappel inférieur à 10%. Le *Sim* a les meilleurs résultats quand le taux de rappel est entre 10% et 30%. Lorsque le nombre des documents restitués augmente, les résultats deviennent quasi-identiques.

Le graphique illustré par la **figure 3.31** montre la grande ressemblance entre ces deux courbes, confirmant le fait que les mesures étudiées sont de même ordre. Néanmoins, les mesures moyennes sont le reflet du comportement général, en y regardant de plus prêt, pour certaines requêtes, le *Sim* majore le *Cosinus*.

Ce graphique montre 20 requêtes (sur 50 traitées) pour lesquelles le *Sim* majore le *Cosinus*. Cela confirme que le *Sim* semble avoir un effet positif dans certains cas.

b. Synthèse

Nous concluons que notre mesure utilisée donne parfois des résultats inférieurs à la mesure témoin (MAP), mais parfois montre un avantage (R-précision). Elle peut donc être envisagée pour classer les documents.

Les mesures utilisées pour évaluer la similarité entre la requête et les documents concernent la qualité de ses réponses. Or, il est intéressant en RI de classer les documents retournés en utilisant un score de pertinence.

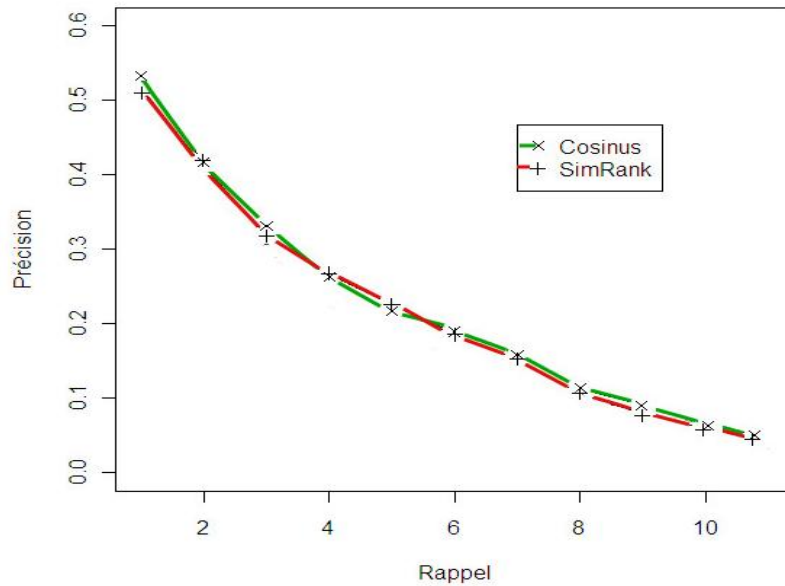


Figure 3.30 : Courbe précision-rappel en 11 points pour Cosinus et Sim

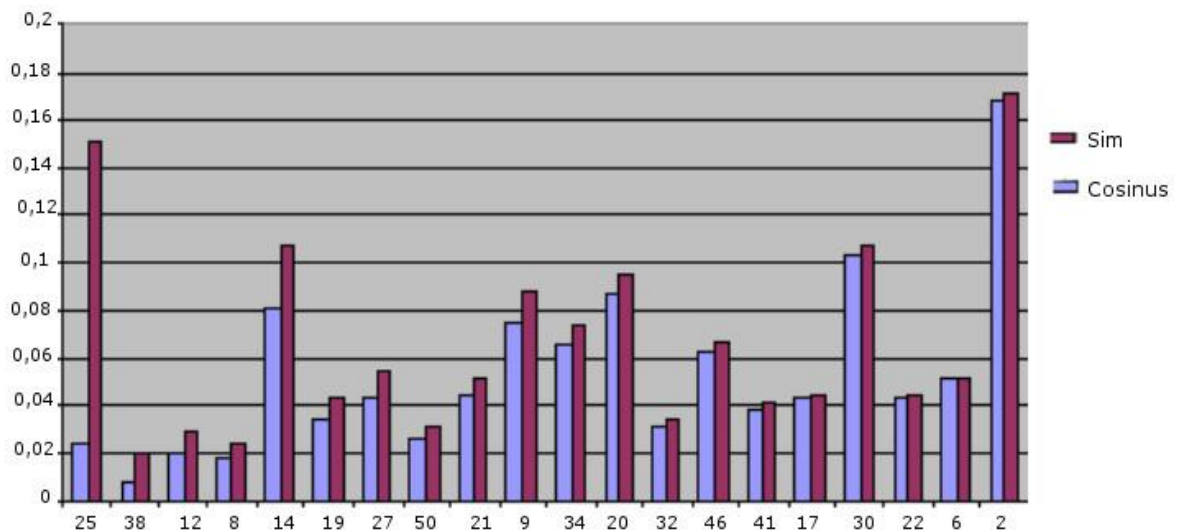


Figure 3.31 : Comparaison Cosinus/Sim des moyennes des précisions n par requête

III.5.2.2. Calcul de pertinence

Afin de déterminer la pertinence du document, nous fixons un seuil de pertinence en se basant sur une mesure donnée. Nous avons choisi la mesure F pour évaluer la liste des documents retournés

Après avoir noté et classé les documents, le nombre de documents à retourner pour obtenir une meilleure mesure F pour chaque requête est calculé. Le seuil expérimental retenu est celui qui maximise la mesure F. Ce nombre de documents à retourner est nommé rang de la meilleure mesure F. Le score obtenu par le document à ce rang est le seuil absolu. Après les études que nous avons faites, nous avons constaté que ce rang est en moyenne égal à 78 avec une variance de 87.

a. Résultats

Les scores Sim varient beaucoup d'une requête à l'autre. Pour observer un éventuel lien entre le nombre de termes de la requête et la hauteur du score du document au rang duquel la meilleure mesure F est obtenue, nous traçons le nuage de points du seuil relatif en fonction du nombre de termes de la requête (cf. **Figure 3.32**).

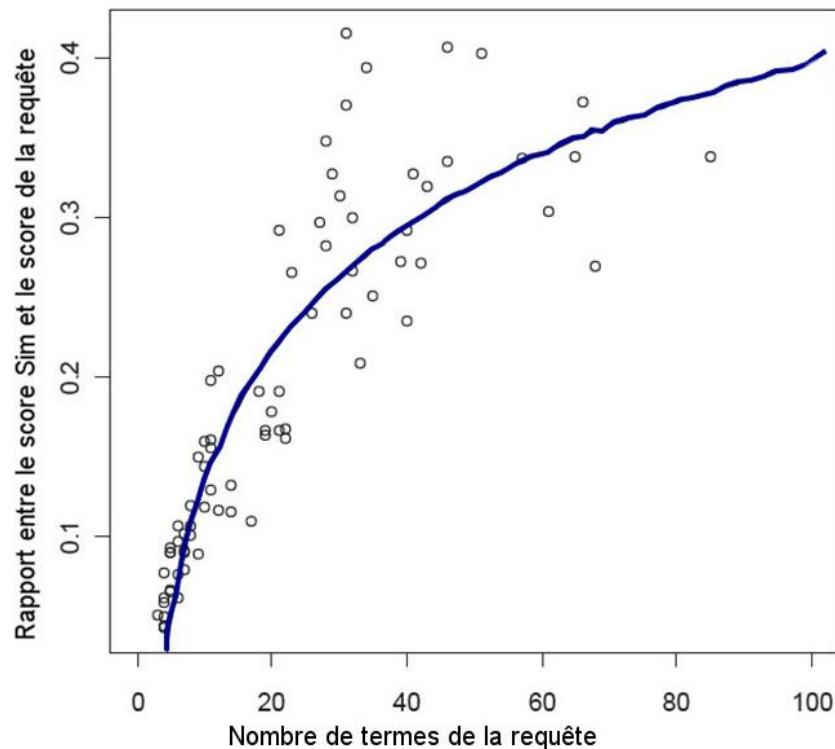


Figure 3. 32 : seuil relatif en fonction du nombre de termes de la requête

b. Synthèse

Le nuage de points suit une courbe de tendance de type logarithmique. Son équation, calculée grâce à la courbe de tendance est :

$$y = 0.11 \ln(x) - 0.113.8$$

avec x = nombre de termes de la requête.

III.5.2.3. Classement

Dans cette partie, nous nous sommes intéressés au jugement de la pertinence des résultats représentés par le nombre de documents retournés. Nous calculons P@1, P@5, P@10, P@15 et P@20. À chaque niveau de pertinence, une note est attribuée entre 0 et 5. 0 correspond à un document totalement inutile ou hors thème. 5 correspond à un document répondant de façon parfaite à la requête. Nous avons demandé l'aide de 50 utilisateurs.

a. Résultats

Tableau 3. 11: Evaluation de la pertinence par jugement de l'utilisateur

Niveau de pertinence	Note
P@1	4.12
P@5	3.98
P@10	3.25
P@15	2.79
P@20	2.21

Nous constatons une baisse générale de pertinence perçue en fonction du nombre de résultats considérés. Ce phénomène est dû à l'efficacité de notre algorithme de classement.

Pour mieux juger la satisfaction de l'utilisateur, nous comparons pour la même série de requêtes, le classement classique des résultats via le classement sémantique guidé par la RE.

Nous avons étudié les 30 premiers résultats à partir d'une série de 50 requêtes dont 20 sont complexes. Ces résultats sont présentés dans le **tableau 3.12**.

Tableau 3. 12: Comparaison de l'efficacité des deux types de classement

Satisfaction de l'utilisateur	Classement sans RE	Classement avec RE
Requêtes simples	8.71	9.14
Requêtes complexes	7.85	8.73
Moyenne	8.28	8.93

b. Synthèse

La comparaison entre le classement classique sans RE via le classement sémantique avec RE a montré que le deuxième type apporte un gain en termes d'efficacité quelque soit le type de requête. Ce gain est dû à l'utilité de l'ajout des nouveaux termes dans la phase de reformulation de la requête.

Conclusion

Dans ce chapitre, nous avons montré l'évaluation des différents modules de notre système.

Le premier module d'analyse est composé de deux étapes : l'extraction des termes simples et l'extraction des termes composés. Nous avons évalué la première étape à l'aide de la mesure de précision et la mesure de rappel. La deuxième étape est évaluée en comparant l'approche proposée à l'approche statistique existante. Nous avons utilisé la précision moyenne et la précision à 5 documents comme mesure d'évaluation.

Pour le module d'indexation, nous avons utilisé le même principe que la deuxième étape du module d'analyse. Cette évaluation nous a permis de constater que notre système donne de meilleur résultat en termes de Rappel/Précision. Ceci peut être dû au fait que nous avons pris en compte la sémantique des concepts en tenant compte de la hiérarchie de concepts dans la RE.

Pour le module de recherche, nous avons comparé les scores obtenus d'une recherche classique à ceux d'une recherche sémantique guidée par la RE. Les résultats obtenus ont montré que la reformulation de la requête procure une amélioration remarquable en termes de pertinence des résultats retournés. Cette amélioration se traduit par un gain en termes de rappel et précision.

Conclusion et perspectives

Au terme de cette thèse, nous nous sommes intéressés aux techniques et aux outils d'analyse, d'indexation et de recherche d'information multimédia. Nous avons étudié le contexte général de l'utilisation de la sémantique pour la représentation de l'information dans les SRI et montré son apport dans l'amélioration de ces trois modules et de leurs performances. Ces modules utilisent dans leurs fonctionnements des ressources sémantiques externes : les ontologies et les thésauri.

Dans le module d'analyse, nous avons extrait les termes simples, composés et spécifiques qui caractérisent un document multimédia. Ensuite, dans le module d'indexation, nous avons dégagé les descripteurs sémantiques qui sont les concepts et les relations sémantiques entre les concepts. L'ensemble des concepts et des relations sémantiques est présenté par la suite sous forme d'un GC pour mieux décrire le contenu d'un document multimédia. Enfin, dans le module de recherche, nous avons proposé une approche de reformulation des requêtes tout en se basant sur les ressources externes et/ou le profil utilisateur.

Dans la suite de cette conclusion, nous donnons nos principales contributions et les perspectives ouvertes par nos travaux.

I. Contribution

I.1. Sur le plan théorique

I.1.1. Module d'analyse

I.1.1.1. Analyse structurelle

- Nous avons proposé des descriptions génériques et indépendamment de type de média. Ces descriptions sont basées sur la fusion des différentes modalités.

L'un des problèmes des méthodes de fusion existantes est la perte potentielle de la corrélation entre les différentes modalités. Par conséquent, nous avons proposé une méthode de fusion pour représenter le contenu audio-visuel tout en tenant compte des aspects multimodaux, temporels et événementiels de la visioconférence. Cette fusion garde la corrélation entre les informations auditives et visuelles.

I.1.1.2. Analyse sémantique

- Nous avons proposé une méthode automatique d'extraction des termes simples. Cette méthode est basée sur l'identification des mots vides. Ces derniers sont déterminés par rapport au domaine d'application. Les mots non vides sont considérés comme des termes simples.
- En utilisant la mesure statistique IMA et en se basant sur la liste des termes simples, nous avons extrait les termes composés. À la différence des travaux

antérieurs, un terme composé peut être composé non seulement par deux ou trois termes simples, mais par plusieurs termes selon le cas.

- Les termes simples et les termes composés sont par la suite pondérés en utilisant les pondérations tf_i et Ctf_i . Uniquement les termes pondérés sont pris en compte.
- La majorité des travaux existants dans la littérature, les termes non pondérés sont ignorés. Mais parfois, nous trouvons un terme de poids faible mais riche du côté sémantique. Ce terme est considéré comme un terme spécifique. Pour bénéficier de l'importance de ces termes dans l'indexation de la visioconférence, nous avons examiné la liste des termes non pondérés pour extraire la liste des termes spécifiques. Cette extraction est semi-automatique.

I.1.2. Module d'indexation

- En utilisant les ressources externes, nous avons transformé les termes en concepts en traitant les ambiguïtés sémantiques existantes. Ainsi, nous avons proposé une démarche de désambiguïsation qui permet de favoriser les concepts ambigus en relation avec un concept non ambigu de la même scène.
- Nous avons, également, proposé une méthode d'extraction des relations sémantiques entre les concepts. Cette méthode est basée sur le principe qu'une relation sémantique n'existe entre deux concepts que si ces deux derniers appartiennent à la même scène et qu'il y'a une relation entre eux dans la ressource externe.

I.1.3. Module de recherche

- Ce module se compose essentiellement de deux processus : le processus de reformulation de la requête initiale de l'utilisateur et le processus d'appariement requête-document.
- En effet, afin de pouvoir distinguer parmi les documents de la collection ceux qui sont pertinents par rapport à une requête de l'utilisateur, il est important d'une part de comprendre les informations structurelles et sémantiques existantes dans un document, et d'autre part, celles de la requête. La première tâche est réalisée dans le module d'analyse et le module d'indexation. La deuxième tâche est faite par le processus de reformulation des requêtes. Contrairement à l'état de l'art, ce processus est basé d'une part sur les ressources externes pour prendre en compte la sémantique des informations existantes dans la requête, et d'autre part, sur les éléments issus du profil utilisateur à la fois.
- Le deuxième processus, existant dans le module de recherche, est le processus d'appariement requêtes-documents. Il consiste à comparer la représentation de la requête avec les représentations des documents. En effet, à chaque saisie d'une requête, le système crée sa représentation sémantique. Puis, il calcule pour chaque couple requête-document, un score qui traduit un degré de pertinence de la distance

des termes de la requête avec ceux des documents. Le système retourne généralement des réponses ordonnées selon le score calculé.

En conclusion, comme les besoins en information, diffèrent d'un utilisateur à un autre, pour qu'un SRI soit efficace et satisfait les besoins de l'utilisateur, il doit prendre en compte toutes les informations existantes dans les documents et concernant l'utilisateur en même temps. Parmi ces informations, nous citons les structures des documents, la sémantique de leurs contenus, leurs contextes d'utilisation et les profils utilisateurs.

I.2. Sur le plan pratique et technique

- Sur le plan technique, nous avons développé un système appelé **O**ntology-based **S**ystem for **S**emantic **V**isio-conference **I**nformation **R**etrieval and **I**ndexation « OSSVIRI ». Ce système est composé de trois modules : analyse, indexation et recherche. Il est basé sur l'approche que nous avons proposée pour chaque module. Il est facilement applicable à d'autres corpus en utilisant d'autres ressources externes. Dans ce système, le module d'analyse est semi-automatique. Le module d'indexation et le module de recherche sont entièrement automatiques.
- Nous avons évalué notre système sur le domaine médical à travers une collection de visioconférences. Malheureusement et l'un des problèmes rencontrés, cette collection ne présente pas un corpus réel et contient uniquement cent visioconférences. Cette limite est due à la non existence des corpus réels des visioconférences contrairement au corpus de documents tels que les collections CLEF, TREC Genomics, etc. et au corpus d'images comme ImageCLEFmed.
- Nous avons testé la performance de notre système en utilisant les métriques issues du domaine de la recherche d'information.
- Durant nos expérimentations et pour valider le module d'indexation, nous avons effectué une comparaison entre les résultats obtenus par notre approche et les résultats obtenus par l'approche linguistique et l'approche statistique existante. Cette comparaison a permis de juger la performance de notre approche et de bien la positionner.
- Nous avons, aussi, testé l'effet d'une variation de la valeur du seuil IMA sur l'efficacité du processus d'extraction des termes composés. Ce test nous a permis de fixer la valeur adéquate.
- Nous avons, également, étudié l'importance de l'utilisation de plusieurs ressources externes dans le processus d'extraction des concepts. Et bien qu'il existe, dans le domaine médical, plusieurs ontologies et thésaurus, la non disponibilité de la majorité de ces ressources est un problème rencontré durant nos travaux.
- Nous avons étudié l'apport de l'utilisation des relations sémantiques dans le module d'indexation. Cette étude montre que l'utilisation des concepts et des relations a un effet positif sur le module de recherche d'information.

- Nous avons expérimenté l'apport de la reformulation des requêtes en utilisant les ressources sémantiques et le profil utilisateur. L'intérêt de cette combinaison répond notamment aux utilisateurs qui ne parviendraient pas à écrire une requête décrivant la totalité de leur besoin d'information. L'efficacité de cette combinaison a été évaluée au travers des expérimentations et comparativement à des résultats sans reformulation, avec reformulation par ressources externes uniquement et avec reformulation par le profil utilisateur uniquement.
- Pour évaluer la performance de notre mesure de similarité *Sim*, nous avons comparé la mesure *Sim* utilisée avec la mesure *Cosinus*. Nous avons constaté que les deux mesures donnent presque les mêmes résultats.
- Pour finir nos expérimentations, nous avons jugé la pertinence des résultats retournés. Nous avons demandé l'aide de 50 utilisateurs pour attribuer des notes aux documents retournés. Nous avons comparé pour la même série de requêtes, les notes attribuées aux documents ordonnés par le classement classique par rapport à celles attribuées aux documents ordonnés par le classement sémantique guidé par la ressource externe. Cette comparaison a montré la performance de notre système.

II. Perspectives

L'évaluation expérimentale de notre contribution a montré son efficacité vis-à-vis des modèles de l'état de l'art, et ouvre des perspectives portant sur l'amélioration : d'extraction des termes spécifiques, la transformation des termes en concepts, la pondération des concepts, l'utilisation du profil utilisateur et la mise à jour de la base d'index.

II.1. Perspectives à court terme

À court terme, nous souhaitons améliorer nos contributions selon les aspects suivants :

- expérimenter notre système sur des collections de documents réels en exploitant d'autres ressources sémantiques qui couvrent bien la totalité du domaine comme OGMS (The Ontology for General Medical Science) ;
- améliorer la pondération des concepts en retenant, avec la cooccurrence d'un concept dans le document, les différents termes qui le composent.

II.2. Perspectives à moyen terme

À moyen terme, nous proposons de générer nos contributions selon deux dimensions, liées aux termes spécifiques et la reformulation des requêtes. Ces perspectives sont :

- améliorer l'extraction des termes spécifiques, en enrichissant, à chaque utilisation de notre système par un expert, les ressources externes en ajoutant ces termes. Pour ce faire, il faudrait développer une nouvelle interface permettant de valider ou de modifier les termes spécifiques afin d'enrichir conceptuellement l'ontologie.
- sauvegarder les reformulations des requêtes traitées afin de les réutiliser dans le cas où l'utilisateur formule la même requête. Cette perspective a pour but de satisfaire l'utilisateur en temps de réponse.

II.3. Perspectives à long terme

À long terme, nous proposons d'améliorer nos contributions selon deux aspects liés au profil utilisateur et la mise à jour de la base d'index. Ces perspectives sont :

- définir un profil utilisateur comportant plusieurs dimensions du contexte dans un système unifié d'accès personnalisé à l'information. Actuellement, seul le centre d'intérêt de l'utilisateur est pris en compte dans la reformulation des requêtes. Le choix des sources et la dimension qualité du profil utilisateur sont nécessaires dans la reformulation.
- mettre à jour les algorithmes de mise à jour des index pour garder la cohérence entre les documents et les index tout en prenant en compte la dynamique des documents et gérer leur impact sur les index. Cette perspective tend à minimiser le délai de mise à jour de l'index.

Nos Publications

Journaux

1. **Ameni Yengui**, Mahmoud Neji : « Information Research in indexed medical video conferences databases », International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume & Issue no: Volume 3, Issue 4, ISSN 2278-6856, pages n°226-232, July - August 2014 (Impact Factor : 3.258).
2. Zwidi A., **YENGUI A.**, NEJI M. : « Research system of semantic information in medical videoconference based on conceptual graphs and domain ontologies », International Journal of Management & Information Technology, (IJMIT) Vol. 7, No. 1, ISSN 2278-5612, Page 969, November, 2013 (Impact Factor : 1.291).

Conférences internationales indexées avec comité de lecture

3. Sameh BACCARI, **Ameni YENGUI**, Mahmoud NEJI : « The use of ontologies for the indexation of medical video-conferencing », IEEE International Conference on Information Technology and e-Services (ICITeS), 24-26 Mars 2012 Sousse, Tunisia.
4. **Ameni Yengui**, Mahmoud Neji : « OSVIRA: Ontology-based System for Semantic Visio-conference Information Retrieval and Annotation », International Business Information Management Conference (14th IBIMA 23-24 June 2010 in Istanbul, Turkey conference proceedings (ISBN: 978-0-9821489-3-8).
5. **Ameni Yengui**, Mahmoud Neji : « Semantic annotation formalism of video-conferencing documents », International Conference IADIS CELDA 2009 Rome, Italy, 20-22 novembre 2009.

Conférences internationales non indexées

6. **Ameni Yengui**, Mahmoud Neji, Rafik Bouaziz : « The Use of Conceptual Graphs to Annotate and to Search Semantically the Pedagogic Videoconferencing », Second Kuwait Conference on e-Services and e-Systems, April 5-7, 2011, Kuwait University
7. Héla ELLEUCH, **Ameni YENGUI**, Mahmoud NEJI : « Audiovisual document modeling by metadata », International International Arab Conference on Information Technology ACIT 2011, Naif Arab University for Security Science (NAUSS), Riyadh, Saudi Arabia, December 11-14, 2011.
8. **Ameni Yengui**, Mahmoud Neji : « A tool of annotation and semantic research of video-conferences founded on Conceptual Graphs », International International Arab

Conference on Information Technology ACIT 2010, Binghazi, Libye (14-16 Décembre 2010).

Articlé accepté et encore non publié

9. **Ameni Yengui**, Mahmoud Neji : « Towards a statistical approach to the analysis, the indexing and the semantic search of medical videoconferences », International Journal of Information Retrieval Research (IJRR), Volume 6.

Bibliographie

A

- [**Abrouk, 06**] : Abrouk L., "*Annotation de documents par le contexte de citation basée sur une ontologie*", Thèse de l'Université des Sciences et Techniques du Languedoc, Soutenue le 27 novembre 2006.
- [**Abbes, 15**] : Abbes R., "*Filtrage et agrégation d'informations vitales relatives à des entités*", Thèse de l'Université Toulouse 3 Paul Sabatier, soutenue le 11/12/2015.
- [**Abderrahim, 13**] : Abderrahim M. A., "*Utilisation Des Ressources Externes Pour la Reformulation des Requêtes Dans un Système de Recherche D'Information*", The Prague Bulletin of Mathematical Linguistics. Volume 99, Issue, Pages 85–97, ISSN (Online) 1804-0462, ISSN (Print) 0032-6585, DOI: 10.2478/pralin-2013-0006, April 2013.
- [**Adbulahhad et al., 11a**] : Adbulahhad K., Chevallet J.-P., Berrut C., "*Exploiting and Extending a Semantic Resource for Conceptual Indexing*" In Troisième Atelier Recherche d'Information SEmantique (RISE 2011), Avignon, France.
- [**Adbulahhad et al., 11b**] : Adbulahhad K., Chevallet J.-P., Berrut C., "*Solving concept mismatch through bayesian framework by extending umls meta-thesaurus*", In CORIA 2011, pp. 311–326.
- [**Ahcene, 05**] : Ahcene B., "*Construction d'une mémoire organisationnelle de formation et evaluation dans un context elearning : le projet MEMORAE*", Thèse de doctorat de Université de Technologie de Compiègne Réalisée dans le cadre du Heudiasyc, 2005.
- [**Airouche, 12**] : AiroucheM., "*Développement de méthodes de segmentation et de suivi d'objets mobiles dans des séquences d'images d'un système multi-caméras*", Thèse de l'Université M'hamed Bougara-Boumerdes, 2012.
- [**Al-Hezmi et al., 06**] : Al-Hezmi A., Rebahi Y., Magedanz T., Arbanowski S., "*Towards an Interactive IPTV for Mobile Subscribers*", In Proceedings of the International Conference on Digital Telecommunications, ICDT apos'06', Côte d'Azur, France, p. 45-45, 2006.
- [**Attia, 13**] : Attia D., "*Segmentation d'images par combinaison adaptative couleur-texture et classification de pixels : Applications à la caractérisation de l'environnement de réception de signaux GNSS*", Ordinateur et société [cs.CY]. Université de Technologie de Belfort-Montbeliard, 2013.
- [**Audeh, 14**] : Audeh B. "*Reformulation sémantique des requêtes pour la recherche d'information ad hoc sur le Web*", Thèse de l'École Nationale Supérieure des Mines de Saint-Étienne, Soutenue le 09 Septembre 2014.

- [Ayache *et al.*, 06] : Ayache S., Gensel J., Quénot G.M., "*Clips Isr experiments at trecvid 2006*", In TRECVID'2006 Workshop, Gaithersburg, MD, USA, 2006.
- [Ayache, 07] : Ayache S. "*Indexation de documents vidéos par concepts par fusion de caractéristiques audio, image et texte*", Thèse de L'INP Grenoble de l'Ecole Doctorale «Mathématiques, Sciences et Technologies de l'Information, le 4 juillet 2007.

B

- [Baccari *et al.*, 12] : Baccari S., Yengui A., Neji M., "*The use of ontologies for the indexation of medical video-conferencing*", ICITeS, Mars, 2012, Sousse, ISBN: 978-1-4673-1167-0.
- [Bazillon, 11] : Bazillon T., "Transcription et traitement de la parole spontanée pour sa reconnaissance automatique", Thèse de LIUM - Laboratoire d'Informatique de l'Université du Maine, 2011.
- [Baziz, 05] : Baziz M., "*Indexation conceptuelle guidée par ontologie pour la recherche d'information*", PhD Thesis, Institut de recherche en informatique de Toulouse, université Paul Sabatier, 2005.
- [Bechhofer *et al.*, 2001] : Bechhofer S., Horrocks I., Goble C., Stevens R., "*OilEd : a Reasonable Ontology Editor for the Semantic Web*", In Proceedings of the Joint German/Austrian Conference on Artificial Intelligence (KI'2001), volume 2174, pp 396–408, Springer-Verlag LNAI, 2001.
- [Béchet *et al.*, 15] : Béchet F., Bendris M., Charlet D., Damnati G., Favre B., Rouvier M., Auguste R., Bigot B., Dufour R., Fredouille C., Linarès G., Senay G., Tirilly P., Martinet J., "*Identification de personnes dans des flux multimédia*", CORIA 2015.
- [Béchet *et al.*, 12] : Béchet F., Auguste R., Ayache S., Charlet D., Damnati G., Favre B., Fredouille C., Levy C., Linares G., Martinet J., "*Percol0 - un système multimodal de détection de personnes dans des documents vidéo*", Actes de la conférence conjointe JEP-TALN-RECITAL 2012, volume 1 : JEP, pages 553–560, Grenoble, 4 au 8 juin 2012.c 2012 ATALA & AFPC.
- [Belkhatir, 05] : Belkhatir M., "*Intégration Signal, Symbole pour l'indexation et la recherche d'images fixes*", Thèse Informatique de l'université Joseph Fourier, Grenoble I, 2005.
- [Ben Abdelali *et al.*, 09] : Ben Abdelali A., Krifa M. N., Touil L., Mtibaa A., Bourenane E., "*Indexation De La Vidéo Par Le Contenu Et Besoin En Architectures Matérielles Dédiées*", SETIT 2009 5th International Conference : Sciences of Electronic, Technologies of Information and Telecommunications March 22-26, 2009 – TUNISIA.
- [Ben Aouicha, 09] : Ben Aouicha M., "*Une approche algébrique pour la recherche d'information structurée*", Thèse de l'Université Paul Sabatier, soutenue le 08 Janvier 2009.

- [Bendris, 11] : Bendris M., "*Indexation audiovisuelle des personnes dans un contexte de télévision*", Thèse de TELECOM ParisTech, soutenue le 7 juillet 2011.
- [Benmokhtar, 09] : Benmokhtar R., "Fusion multi-niveau pour l'indexation et la recherche multimédia par le contenu sémantique", Thèse de doctorat d'Ecole Doctorale d'Informatique, Télécommunications et Électronique de Paris (EDITE), 2009.
- [Benmokhtar, 10] : Benmokhtar R., "*Fusion multi-niveaux pour l'indexation et la recherche multimédia par le contenu sémantique*", Thèse de doctorat de l'Université ParisTech, 2010.
- [Benosman, 15] : Benosman A., "*Médiation sémantique de données basée sur les services web : Application dans le domaine médical*", Thèse de l'Université Abou Bekr Belkaid Tlemcen, Soutenue publiquement le 21/11/2015.
- [Bermejo *et al.*, 11] : Bermejo Nievas E., Deniz Suarez O., Bueno García G., Sukthakar R., "*Violence Detection in Video Using Computer Vision Techniques*" In : *Computer Analysis of Images and Patterns*, pp. 332–339, Springer Berlin Heidelberg, 2011.
- [Bertrand, 06] : Bertrand D., "*Modèles d'indexation multimédia pour la description automatique de films de cinéma*", Thèse de doctorat de l'Université de Paris VI-Pierre et Marie Curie, 2006.
- [Bhatnagar *et al.*, 13] : Bhatnagar G., Wu Q.M.J., Zheng L., "*Directive Contrast Based Multimodal Medical Image Fusion in NSCT Domain*" *Multimedia, IEEE Transactions on* (Volume:15 , Issue: 5), Page(s): 1014 – 1024, ISSN : 1520-9210.
- [Bhushan *et al.*, 97] : Bhushan N., Rao A.R., Lohse G.L., "*The texture lexicon : understanding the categorization of visual texture terms and their relationship to texture image*", *Cognitive Science* 21(2): 219-246 (1997).
- [Bigot, 11] : Bigot B., "*Recherche du rôle des intervenants et de leurs interactions pour la structuration de documents audiovisuels*", Thèse de doctorat de l'Université Toulouse III Paul Sabatier (UT3 Paul Sabatier), juillet 2011.
- [Blair, 90] : Blair DC., "*Language and representation in information retrieval*", In *Proceedings of Elsevier Science Publications*, Amsterdam, 1990.
- [Bost *et al.*, 15] : Bost X., Linarès G., "*Détection de locuteurs dans les séries TV*", CORIA, 2015 : 269-284.
- [Boubekeur *et al.*, 07] : Boubekeur F., Tamine-Lechani L., "*Recherche d'Information Flexible Basée CP-Nets*", CORIA 2007.
- [Bougares, 12] : Bougares F., "*Attelage de systèmes de transcription automatique de la parole*", Thèse de l'Université du Maine, 2012.
- [Boughanem *et al.*, 07] : Boughanem M., Said L'hadj L., Amrouche K., "*Vers un modèle de langue mixte concepts-mots pour la recherche d'information*", Conférence en Recherche d'Informations et Applications - CORIA 2007, 4th French Information Retrieval Conference, Saint-Étienne, France, March 28-30, 2007.

- [**Bouramoul, 11**] : Bouramoul A., "*Recherche d'information contextuelle et sémantique sur le web*", Thèse de l'Université Mentouri de Constantine Faculté des Sciences de l'Ingénieur, Département d'Informatique, 2011.
- [**Bouzir et al., 16**] : Bouzir A., Latiri C., Gaussier E., "*Expansion de requêtes par apprentissage*", CORIA-CIFED 2016 : 267-282.

C

- [**Camus, 11**] : Camus M., "*Identification audio pour la reconnaissance de la parole*", Thèse de l'Université Paris Descartes soutenue le 30 novembre 2011.
- [**Carlioni, 08**] : Carlioni O., "*Introduction de raisonnement dans un outil industriel de gestion des connaissances*", Thèse de l'Université Montpellier II Soutenue le 24 novembre 2008.
- [**Charhad, 05**] : Charhad M., "*Modèles de documents vidéo basés sur le formalisme des graphes conceptuels pour l'indexation et la recherche par le contenu sémantique*", Thèse de doctorat de l'Université JOSEPH FOURIER- Grenoble I, 2005.
- [**Charnois, 11**] : Charnois T., "*Accès à l'information : vers une hybridation fouille de données et traitement automatique des langues*", HDR de l'Université de Caen, soutenue le 1er décembre 2011.
- [**Charlet et al., 09**] : Charlet J., Vandenbussche P.-Y., "*Méta-modèle général de description de ressources terminologiques et ontologiques*", IC 2009 – 20èmes Journées Francophones d'Ingénierie des Connaissances, May 2009, Hammamet, Tunisie
- [**Chebil et al., 15**] : Chebil W., Soualima L.F., Omri M.N., Darmori S.I., "*Biomedical Concepts Extraction Based on Possibilitic Network and Vector Space Model*", 15th conference on Artificial Intelligence in Medecine.
- [**Chein et al., 07**] : Chein M., MU Mugnier M.L., "*Positive Nested Conceptual Graphs*", In Proc. ICCS'97, LNAI, p. 95–109 : Springer, 2007.
- [**Chesnais et al., 12**] : Chesnais T., Allezard N., Dhome Y., Chateau T., "*Création automatique d'un détecteur adapté à la scène*", RFIA 2012 (Reconnaissance des Formes et Intelligence Artificielle), Jan 2012, Lyon, France. pp.978-2-9539515-2-3, 2012.
- [**Chifu, 15**] : Chifu A. G., "*Adaptation des systèmes de recherche d'information aux contextes : le cas des requêtes difficiles*", Thèse de L'Université de Toulouse, soutenue le 15/06/2015.
- [**Claude, 12**] : Claude G., "*Modélisation de documents et recherche de points communs Proposition d'un framework de gestion de fiches d'anomalie pour faciliter les maintenances corrective et préventive*", Thèse de l'université de Toulouse, soutenue le 16 mai 2012.

- [Claveau *et al.*, 04] : Claveau V., Sébillot P., "*Extension de requêtes par lien sémantique nom-verbe acquis sur corpus*", In Actes de la 11ème conférence de Traitement automatique des langues naturelles, Fès, Maroc, avril 2004.
- [Codreanu, 10] : Codreanu D., "*Spécification et utilisation d'un modèle générique de description d'algorithmes d'indexation multimédia*", Master de Recherche en Informatique de l'université de Toulouse, soutenu le 23 Juin 2010.
- [Cohen *et al.*, 10] : Cohen A., Attia D., Meurie C., Ruichek Y., "*Une méthode de segmentation hybride par combinaison adaptative des informations texture et couleur*", MajecSTIC2010, Bordeaux, France, du 13 au 15 octobre 2010.

D

- [Drame, 14] : Drame K., "*Contribution à la construction d'ontologies et à la recherche d'information : application au domaine médical*", Thèse de L'Université de Bordeaux, Soutenue le 10 décembre 2014.
- [Davis *et al.*, 06] : Davis M., Smith M., Stentiford F., Bambidele A., Canny J., Good N., King S., Janakiraman R., "*Using context and similarity for face and location identification*", In Proceedings of the IS& T/SPIE 18th Annual Symposium on Electronic Imaging Science and Technology Internet Imaging VII, page cdrom, San Jose, California, 2006. IS& T/SPIE Press.
- [Delezoide, 06] : Delezoide B., "*Modèles d'indexation multimédia pour la description automatique de films de cinéma*", Thèse de l'Université Paris VI, soutenue le 24 Avril 2006.
- [Derbas, 14] : Derbas N., "*Contributions à la détection de concepts et D'événements dans les documents vidéos*", Thèse DE L'Université de Grenoble, soutenue le 30 Septembre 2014.
- [Dessein *et al.*, 11] : Dessein A., Arshia C., "*Segmentation statistique de flux audio en temps-réel dans le cadre de la géométrie de l'information*", GRETSI - 23e Colloque du Groupe de Recherche et d'Etudes du Traitement du Signal, Sep 2011, Bordeaux, France. 2011.
- [Djamel *et al.*, 08] : Djamel A., Zighed A., Venturini G., "*Extraction des connaissances : Etat et perspective*", Revue des Nouvelles Technologies de l'Information Sous la direction de Djamel A. Zighed et Gilles Venturini RNTI-E-5 2008.
- [Dinh *et al.*, 10] : Dinh D., Tamine L., "*Vers un modèle d'indexation sémantique adapté aux dossiers médicaux de patients*", Conférence francophone en Recherche d'Information et Applications, CORIA 2010, 2010, pp.325-336.
- [Dinh, 2012] : Dinh B.D., "*Accès à l'information biomédicale : vers une approche d'indexation et de recherche d'information conceptuelle basée sur la fusion de ressources termino-ontologiques*", Thèse de l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier), septembre 2012.

- [Dufour, 10] : Dufour R., "*Transcription automatique de la parole spontanée*", Thèse du Laboratoire d'Informatique de l'Université du Maine (LIUM), 01/12/2010.
- [Dumont *et al.*, 12] : Dumont E., Quénot G., "*A local temporal context-based approach for TV news story segmentation*". In : Multimedia and Expo (ICME), 2012 IEEE International Conference on, pp. 973–978, IEEE, 2012.

E

- [Ercolessi *et al.*, 12] : Ercolessi P., Sénac C., Bredin H., Mouysset S., "*Vers un résumé automatique de séries télévisées basé sur une recherche multimodale d'histoires*", Document numérique, 2012/2 (Vol. 15), Pages : 124, ISBN : 9782746239760.
- [Ercolessi, 13] : Ercolessi P., "*Extraction multimodale de la structure narrative des épisodes de séries télévisées*", Thèse de l'Université Toulouse 3 Paul Sabatier, soutenue le 26/09/2013.

F

- [Farabet *et al.*, 13] : Farabet C., Couprie C., Najman L., LeCun Y., "*Learning hierarchical features for scene labeling*". Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 35, No. 8, pp. 1915–1929, 2013.
- [Feng *et al.*, 12] : Feng B., Ding P., Chen J., Bai J., Xu S., Xu B., "*Multi-modal information fusion for news story segmentation in broadcast video*". In : Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, pp. 1417–1420, IEEE, 2012.
- [Ferecatu *et al.*, 05] : Ferecatu M., Boujemaa M., Crucianu M., "*Hybrid visual and conceptual image representation within active relevance feedback context*", In 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, page 209-216, Singapore, November 2005.
- [Fensel, 00] : Fensel D., "*Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*", Springer, 2000.
- [Fernandez-Luna *et al.*, 04] : Fernandez-Luna J., Compos L., Huete J., "*Using context information in structured document retrieval : An approach based on impudence diagrams*", Information Processing and Management, 40 : 829,847, 2004.
- [Florea *et al.*, 05] : Florea F., Rogazan A., Benschair A., Dacher J.N., Darmoni S., "*Modality categorization by textual annotations interpretation in medical imaging*", In connecting Medical Informatics and Bio-Informatics, Medical Informatics Europe, page 1270-1275, R.E. et al, 2005.
- [Fuhr *et al.*, 03] : Fuhr N., Grossjohann K., "*XIRQL : a query language for information retrieval in XML documents*", In Proceedings of SIGIR 2001, Toronto, Canada, 2003.

G

- [Gagnon, 13] : Gagnon O., " *Indexation de documents web à l'aide d'ontologies*", mémoire présenté en vue de l'obtention du diplôme de maîtrise ès sciences appliquées, école polytechnique de MONTRÉAL, 2013.
- [Gandon, 08] : Gandon F. L., " *Graphes RDF et leur Manipulation pour la Gestion de Connaissances*", Mémoire d'Habilitation à Diriger les Recherches de l'Université de Nice, 5 novembre 2008.
- [Geoffroy, 13] : Geoffroy P., " *Indexation automatique de contenus audio musicaux*", Mémoire d'Habilitation à Diriger des Recherches Université Pierre et Marie Curie (Université Paris VI), Soutenue le 12 avril 2013.
- [Geva, 04] : Geva S., " *Gpx-gardens point xml information retrieval at inex 2004*", In INEX 2004 Workshop Proceedings, pages 211,223. Dagstuhl, Germany, December 2004.
- [Ghaoui *et al.*, 05] : Ghaoui C., Bannore V., Jain L.C., Jain M., " *Knowledge-Based Virtual Education : User-Centred Paradigms*", Springer, 2005.
- [Gruber, 93] : Gruber T.R., " *A translation approach to portable ontology specifications*", Knowl. Acquis, 5(2):199–220, 1993.
- [Guarino, 99] : Guarino N., " *Some organizing principles for a unified top-level ontology*", In Proceedings of AAAI Spring Symposium on Ontological Engineering., pages 1.1–1.15, Stanford, 1999.
- [Guyot, 11] : Guyot B., " *Introduction à l'ingénierie documentaire et aux sciences de l'information*", Institut des sciences et techniques de la documentation, Octobre 2011.

H

- [Habibian *et al.*, 14] : Habibian A., Mensink T., Snoek C.G., " *Composite Concept Discovery for Zero-Shot Video Event Detection*", ICMR'14, April 1–4, 2014, Glasgow, United Kingdom 2014.
- [Hamadi, 14] : Hamadi A., " *Utilisation du contexte pour l'indexation sémantique des images et vidéos*", Thèse de l'Université de Grenoble soutenue le 23 Octobre 2014.
- [Hamadi *et al.*, 15] : Hamadi A., Mulhem P., Quénot G., " *Extended conceptual feedback for semantic multimedia indexing*", in the International Journal Multimedia Tools and Applications, volume 74, Issue 4, pp 1225-1248, February 2015.
- [Harrathi *et al.*, 05] : Harrathi F., Roussey C., Maisonnasse L., Calabretto S., Gammoudi M. M., " *Indexation sémantique des documents multilingues*", liris.cnrs.fr, 2005.
- [Harrathi, 09] : Harrathi F., " *Extraction de concepts et de relations entre les concepts à partir des documents multilingues : Approche statistique et ontologique*", Thèse de l'Institut National des Sciences Appliquées de Lyon, 2009.

- [**Harrathi et al., 11**] : Harrathi F., Maisonnasse L., Roussey C., Calabretto S., " *Approche statistique versus approche linguistique pour l'indexation sémantique des documents multilingues*", Document numérique 08/2011 ; 14(2):193-204. DOI:10.3166/dn.14.2.193-204.
- [**Hjelsvold et al., 94**] : Hjelsvold R., Midtstraum R., " *Modeling and Querying Video Data*", proceedings of the 20th VLDB Conference Santiago, Chile, 1994.
- [**Huart, 07**] : Huart J., " *Extraction et analyse d'objets-clés pour la structuration d'images et de vidéos*", Thèse de l'Institut National Polytechnique de Grenoble, soutenue le 14 février 2007.
- [**Hubert, 05**] : Hubert G., " *A voting method for XML retrieval*", In Norbert Fuhr, Mounia Lalmas, and Saadia Malik, editors, *Advances in XML Information Retrieval : Third International Workshop of the Initiative for the Evaluation of XML Retrieval, INEX 2004, LNCS 3493 / 2005*, Dagstuhl, Germany, , pages 183{196. Springer-Verlag GmbH, mai 2005. Dates de conférence : mai 2005.

I

- [**Ingwersen, 94**] : Ingwersen P., " *Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction*", In Proceedings of the Seventeenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval., pages 101-110, 1994.
- [**Isaac, 05**] : Isaac A., " *Conception et utilisation d'ontologies pour l'indexation de documents audiovisuels*", Thèse de l'Université de Paris IV – SORBONNE, 2005.

J

- [**Jonquet et al., 09**] : Jonquet C., Shah N., Musen M.A., " *Un service web pour l'annotation sémantique de données biomédicales avec des ontologies*", 13^{ème} Journées Francophones d'Informatique Médicale, JFIM'09, Nice, France, 2009.

K

- [**Kakade et al., 05**] : Kakade V., Raghavan P., " *Encoding XML in vector spaces*", In Proceedings of ECIR 2005, Saint Jacques de Compostelle, Spain, 2005.
- [**Kassel, 09**] : Kassel G., " *Vers une ontologie formelle des artefacts*", Actes des 20es Journées Francophones d'Ingénierie des Connaissances, IC 2009.
- [**Kharbouche, 06**] : KHARBOUCHE S., " *Fonctions de croyance et indexation Multimodale*", Thèse de doctorat en informatique, Université de Rouen, Soutenue le 8 décembre 2006.

- [**Kim et al., 05**] : Kim C. S., Kim D. Y., Bae T. M., Ro Y. M., "*Versatile video on demand system*", In IConsumer Electronics (ISCE 2005), Proceedings of the Ninth International Symposium on Volume, Issue 19, p. 441-444, 2005.
- [**Kompaoré, 08**] : Kompaoré N.D.Y., "*Fusion de systèmes et analyse des caractéristiques linguistiques des requêtes : vers un processus de RI adaptatif*", Thèse de doctorat en informatique, Université Paul Sabatier de Toulouse, 2008.
- [**Kostadinov, 08**] : Kostadinov D., "*Personnalisation de l'information : une approche de gestion de profils et de reformulation de requêtes*", Thèse de L'Université de Versailles SAINT-QUENTIN-EN-YVELINES, 2008.
- [**Khoury et al., 14**] : Khoury E., Sénac C., Joly P., "*Audiovisual diarization of people in video content*". Multimedia Tools Appl., Vol. 68, No. 3, pp. 747–775, 2014.
- [**Krizhevsky et al., 12**] : Krizhevsky A., Sutskever I., Hinton G. E., "*Magenet classification with deep convolutional neural networks*". In : Advances in neural information processing systems, pp. 1097–1105, 2012.
- [**Krovetz, 97**] : Krovetz R., "*Homonymy and polysemy in Information Retrieval*", In Proceedings of ACL/EACL'97.

L

- [**Lajmi, 11**] : Lajmi S., "*Annotation et recherche contextuelle des documents multimédias socio-personnels*", Thèse de l'Université de Lyon, Soutenue le 11 mars 2011.
- [**Lame, 02**] : Lame G., "*Construction d'ontologie a partir de textes. Une ontologie du droit dédiée à la recherche d'information sur le Web*", PhD thesis, Ecole des Mines de Paris, 2002.
- [**Latour, 14**] : Latour M., "*Du besoin d'informations à la formulation des requêtes : étude des usages de différents types d'utilisateurs visant l'amélioration d'un système de recherche d'informations*", Thèse de l'Université de Grenoble, soutenue publiquement le 24 juin 2014.
- [**Le Maguer et al., 15**] : Le Maguer S., Haman t., Grabar N., Claveau V., "*Recherche d'information médicale pour le patient Impact de ressources terminologiques*", Conférence en Recherche d'Information et Application, CORIA 2015.
- [**LeCu et al., 10**] : LeCun Y., Kavukcuoglu K., Farabet C., "*Convolutional networks and applications in vision*". In : Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pp. 253–256, IEEE, 2010.
- [**Lee et al., 99**] : Lee J.C.M., Li Q., Xiong W., "*Automatic and dynamic video manipulation*", in Furthi'99, pp 317-343, 1999.
- [**Leung et al., 12**] : Leung K.W.-T., Lee D.L., Ng W., Fung, H.Y. "*A Framework for Personalizing Web Search with Concept-based User Profiles*", ACM Transactions on Internet Technology (TOIT), 11(4) :1–29, 2012.

[Lu *et al.*, 10] : Lu M.M., Xie L., Fu Z.H., Jiang D.M., Zhang Y.N.. "Multimodal feature integration for story boundary detection in broadcast news", In : Chinese Spoken Language Processing (ISCSLP), 2010 7th International Symposium on, pp. 420–425, IEEE, 2010.

M

[Maisonasse, 08] : Maisonasse L., "les supports de vocabulaires pour les systèmes de recherche d'information orientés précision : application aux graphes pour la recherche d'information médicale". Université Joseph Fourier-Grenoble.I. Thèse de doctorat en Informatique, 2008.

[Maisonasse *et al.*, 09] : Maisonasse L., Gaussier E., Chevallier J.P., "Combinaison d'analyses sémantiques pour la recherche d'information médicale", Atelier Recherche d'Information Sémantique RISE, Toulouse 26 mai 2009, Associé au 27^{ème} Congrès INFORS'I 2009.

[Mallek., 11] : Mallek I., "De nouveaux facteurs pour l'exploitation de la sémantique d'un texte en Recherche d'Information", Thèse de l'Université de Toulouse soutenue le 11 Juillet 2011.

[Manser, 12] : Manser M., "État de l'art sur l'acquisition de relations sémantiques entre termes : contextualisation des relations de synonymie", Actes de la conférence conjointe JEP-TALN-RECITAL 2012, volume 3 : RECITAL, pages 163–175, Grenoble, 4 au 8 juin 2012. c 2012 ATALA & AFCEP.

[Martinet, 04] : Martinet J., "Un modèle vectoriel relationnel de recherche d'information adapté aux images", Thèse de l'Université de Grenoble 1, 2004.

[Mass *et al.*, 04] : Mass Y., Mandelbrod M., "Component ranking and automatic query refinement for XML retrieval", In INEX 2004 Workshop Proceedings, pages 73, 84. Dagstuhl, Germany, December 2004.

[Mataoui, 15] : Mataoui, M., "Reformulation de requêtes : Application aux systèmes de recherche d'information dans les documents XML", ouvrage de l'Édition Universitaire Européenne, 2015.

[Mazloom *et al.*, 13] : Mazloom M., Gavves E., Van de Sande K., Snoek C., "Searching informative concept banks for video event detection". In : Proceedings of the 3rd ACM conference on International conference on multimedia retrieval, pp. 255–262, ACM, 2013.

[Mechkour, 95] : Mechkour M., "EMIR², un modèle étendu de représentation et de correspondance d'images pour la recherche d'informations. Application à un corpus d'images historiques", Thèse Informatique de l'Université Joseph Fourier, Grenoble I, novembre 1995.

[Merzougui *et al.*, 11] : Merzougui G., Djoudi M., Behaz A., "Conception and Use of Ontologies for Indexing and Searching by Semantic Contents of Video Courses",

IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 6, No 3, November 2011 ISSN (Online): 1694-0814.

- [**Mothe, 11**] : Mothe J., "*Recherche d'information contextuelle : le cas des requêtes*", In : Recherche d'information et modélisation usagers. In P. Bellot (Ed.) Recherche d'information contextuelle, assistée et personnalisée. Paris : Hermès, 2011.
- [**Moussallam, 12**] : Moussallam M., "*Représentations Redondantes et Hiérarchiques pour l'Archivage et la Compression de Scènes Sonores*", Thèse de TELECOM ParisTech soutenue le e 18 décembre 2012.
- [**Mustafa, 05**] : Mustafa E. W. "*Indexation humaine et indexation automatisée : la place du terme et de son environnement*", [http://perso.univ-lyon2.fr/~thoiron/ JS%20LTT%202005/pdf/Mustafa.pdf](http://perso.univ-lyon2.fr/~thoiron/JS%20LTT%202005/pdf/Mustafa.pdf).

N

- [**Neches et al., 91**] : Neches R., Fikes R., Finin T., Gruber T., Patil R., Senator T., Swartout W.R., "*Enabling Technology For Knowledge Sharing*", AI Magazine, 1991.
- [**Nedjar, 14**] : Nedjar I., "*CMBIR (Content Medical Based Image Retrieval) développement d'outil logiciel d'annotation d'images médicales, utilisant les méthodes d'indexation par descripteurs invariants de contenus*", Mémoire de magister de l'Université Abou Bakr Belkaid– Tlemcen, 2014.
- [**Niang et al., 13**] : Niang C., Bouchou B., LO M., SAM Y., "*Ré-écriture de requêtes dans un système d'intégration sémantique*", In : Proceedings EGC 2013 Toulouse, 2013, pp.383-388.

O

- [**Ounis et al., 98**] : Ounis I., Pasca M., "*Relief : Combining expressiveness and rapidity into a single system*", in 21st International ACM SIGIR, ACM Press, Melbourne, Australia, August 24-28, pp 266-274, 1998.
- [**Over et al., 05**] : Over P., Karray W., Smeaton A., "*TRECVID 2005- An Introduction*", TRECVID 20056 Text REtrieval Conference TRECVID Workshop, 2005.

P

- [**Pal et al., 11**] : Pal, A., Counts, S. "*Identifying topical authorities in microblogs*", In Proceedings of the Conference on Web Search and Data Mining, WSDM '11, pages 45–54. ACM.

- [**Parvaix, 10**] : Parvaix M., "*Séparation de sources audio informée par tatouage pour mélanges linéaires instantanés stationnaires*", Thèse de l'Université de Grenoble, soutenue le 23 Septembre 2010.
- [**Philipp-Foliguet et al., 06**] : Philipp-Foliguet S., Logerot G., Constant P., Gosselin P.H., Lahanier C., "*Multimedia indexing and fast retrieval based on a vote system*", In IEEE International Conference on Multimedia and Expo, page 1781-1784, Toronto, Ontario, Canada, July 2006.
- [**Piwowarski et al., 02**] : Piwowarski B., Faure G., Gallinari P., "*Bayesian networks and inex*", In Proceedings of the First Workshop of the INiative for the Evaluation of XML Retrieval (INEX), pages 149,154. Dagstuhl, Germany, December 2002.
- [**Plu et al., 11**] : Plu M., Heinecke J., "*Interprétation linguistique de requêtes pour un moteur de questions réponses grand public*". In : Actes CORIA'2011, [8 p.].
- [**Poignant, 11**] : Poignant J., "*Détection et reconnaissance de texte dans les documents vidéos - Et leurs apports à la reconnaissance de personnes*", Gabriella Pasi, Patrice Bellot. CORIA 2011 - Conférence en Recherche d'Information et Applications - 6e Rencontres Jeunes Chercheurs en Recherche d'Information (RJCRD), Mar 2011, Avignon, France. Editions Universitaires d'Avignon, pp.409-414, 2011.
- [**Ponte et al., 98**] : Ponte J.M., Croft W.B., "A language modelling approach to information retrieval", In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 40{48, 1998.
- [**Psyche et al., 03**] : Psyche V., Mendes O., Bourdeau J., "*Apport de l'ingénierie ontologique aux environnements de formation à distance*", volume 10 of Article de recherche, Sticef : Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation, 2003.

Q

- [**Qiu et al., 93**] : Qiu Y., Frei H.P., "*Concept Based Query Expansion*", In Proceedings of the 16th ACM SIGIR Conference on Research and Development in Information Retrieval, 160-169, Pittsburg, USA 1993.

R

- [**Radinsky et al., 13**] : Radinsky K., Svore K.M., Dumais S.T., Shokouhi M., Teevan J., Bocharov A., Horvitz, E., "*Behavioral Dynamics on the Web : Learning, Modeling, and Prediction*", ACM Transactions on Information Systems (TOIS), 31(3) :1-37, 2013.
- [**Ralalason, 10**] : Ralalason B., "*Représentation multi-facette des documents pour leur accès sémantique*", Thèse de l'Université Toulouse III - Paul Sabatier, 2010.

- [**Raimbault et al., 08**] : Raimbault T., Briand H., Genest D.,Lehn R.,Loiseau S., "*Une synthèse des modèles de représentation des connaissances à base de Graphes Conceptuels et OWL*", Revue des nouvelles technologies de l'information, n° 12, pp. 45-66.
- [**Rastier, 04**] : Rastier F., "*Ontologie(s)*", Revue des sciences et technologies de l'information, 18(1) :15–40, 2004.
- [**Ribeiro et al., 13**] : Ribeiro C., Mucheroni M. L., "*Dynamic Indexation in Video Metadata*", Procedia - Social and Behavioral Sciences Volume 73, 27 February 2013, Pages 551–555.
- [**Rombaut et al., 10**] : Rombaut M., Simac-Lejeune A., Lambert P., "*Points d'intérêt spatio-temporels pour la détection de mouvements dans les vidéos*", MajecSTIC'2010, Oct 2010, Bordeaux, France. pp.n.c., 2010.
- [**Roy, 11**] : Roy A., "*Boosting Localized Features for Speaker and Speech Recognition*", Thèse de Ecole Polytechnique FEDERALE de Lausanne, Suisse, 11 Novembre 2011.

S

- [**Said L'hadj, 09**] : Said L'hadj L., "*Recherche conceptuelle d'information modèle d'indexation mixte : concepts-mots*", Magister en Informatique de l'Ecole nationale Supérieure d'Informatique, Algérie 2009.
- [**Salton, 71**] : Salton G., "*A comparaison between manual and automatic indexing methods*", In Proceedings of Journal of American documentation, 1971.
- [**Salton, 86**] : Salton G., "*Another look at automatic text-retrieval systems*", Commun. ACM, 29(7) : 648–656, 1986.
- [**Salton, 89**] : Salton G., "*Automatic text processing : The transformation, analysis and retrieval of information by computer*", Addison-Wesley publishing, MA, 1989.
- [**Sargent et al., 10**] : Sargent G., Bimbot F., Vincent E., "*Un système de détection de rupture de timbre pour la description de la structure des morceaux de musique*", publié dans "Journées d'Informatique Musicale 2010 (2010)".
- [**Saulier et al., 14**] : Soulier L., Tamine L., Bahsoun W., "*Un modèle de recherche d'information collaborative basé sur l'expertise des utilisateurs*". Conférence en Recherche d'Information et Applications 2014 (CORIA 2014), Mar 2014, Nancy, France. pp.35-50, 2014.
- [**Schwab, 05**] : Schwab D., "*Approche hybride - lexicale et thématique - pour la modélisation, la détection et l'exploitation des fonctions lexicales en vue de l'analyse sémantique de texte*", Thèse de l'Université des Sciences et Techniques du Languedoc,2005.
- [**Shafer, 76**] : Shafer G., "*A mathematical theory of evidence*", Princeton, NJ : Princeton University Press, 1976.
- [**Smadja, 93**] : Smadja F., "*Retrieving collocations from text: Xtract*", Computational linguistics, 1993.

- [Soulie *et al.*, 13] : Soulier, L., Tamine, L., and Bahsoun, W., "A Collaborative Document Ranking Model for a Multi-faceted Search", In Proceedings of the Asia Information Retrieval Societies Conference, AIRS '13, pages 109–120. Springer, 2013.
- [Sowa, 84]: Sowa J.F., "Conceptual structures: information processing in mind and machine", Addison-Wesley, 1984.
- [Sparck-Jones *et al.*, 72] : Sparck-Jones K., Needham R., "Automatic theme classification and retrieval", Information Processing and Management, 4 : 91,100, 1972.
- [Sy *et al.*, 12] : Sy M-F., Ranwez S., Montmain J., Ranwez V., "OBIRS-feedback une méthode de reformulation utilisant une ontologie de domaine", CORIA 2012, pp 135-150, Borderaux, 21-23 mars 2012.

T

- [Tambellini, 07] : Tambellini C., "Un système de recherche d'information adapté aux données incertaines : adaptation du modèle de langue", Thèse de doctorat en informatique, Université de Nice-Sophia Antipolis-UFR sciences, 2007.
- [Tamine, 00] : Tamine L., "Optimisation de requêtes dans un système de recherche d'information approche basée sur l'exploitation de techniques avancées de l'algorithmique génétique". Thèse de l'Université Paul Sabatier, Décembre 2000.
- [Tamine *et al.*, 03] : Tamine L., Chrisment C., Boughanem M., " Multiple query evaluation based on an enhanced genetic algorithm", Information Processing & Management, 39(2): 215. 231, 2003.
- [Tapu, 12] : TAPU R.G., "Segmentation et structuration des documents vidéos pour l'indexation". PhD thesis, Telecom SUDPARIS et l'université Pierre et Marie Curie, Soutenue le 07 Décembre 2012.
- [Tchuenta, 13] : Tchuenta D., "Modélisation et dérivation de profils utilisateurs à partir de réseaux sociaux : approche à partir de communautés de réseaux k-égocentriques", Thèse l'Université Toulouse III - Paul Sabatier, Janvier 2013.
- [Texier, 05] : Texier R., "Taxinomies, thésaurus et ontologies", EliKya, intelligence des organisations, 2005.
- [Thiault, 11] : Thiault F., " Communauté de pratique et circulation des savoirs : la communauté des enseignants documentalistes membres de la liste de discussion Cdidoc", Thèse de l'Université Charles de Gaulle - Lille III, Soutenance le 14 novembre 2011.
- [Tilley *et al.*, 04]: Tilley C.B., Willis J., "Unified Medical Language System Basics", National Library of Medicine, 1 November 2004.

[Tortman *et al.*, 03] : Tortman A., O'Keefe R. A., "*Identifying and ranking relevant document element*". In Proceedings of INEX 2003 Workshop, pages 149,154. Dagstuhl, Germany, December 2003.

[Turtle *et al.*, 90] : Turtle H., Croft W.B., "*Inference networks for document retrieval*", In A. Bookstein, Y. Chiarmella, G. Salton, and V. Raghavan, editors, Proceedings of ACM SIGIR, pages 1,24, 1990.

V

[Van De Sande *et al.*, 10] : Van De Sande K. E., Gevers T., Snoek C. G., "*Evaluating color descriptors for object and scene recognition*". Pattern Analysis and Machine Intelligence, IEEE.

[Vincent, 12] : Vincent E., "*Contributions to audio source separation and content description. Signal and Image processing*", HDR de l'Université Rennes 1, le 23/11/2012.

[Vinet, 13] : Vinet H., "*Traitement par le contenu des signaux musicaux*", Article dans des revues "E-dossiers de l'audiovisuel", 2013, pp.1-1.

W

[Wang *et al.*, 11] : Wang H., Klaser A., Schmid C., Liu C.-L., "*Action recognition by dense trajectories*". In : Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp. 3169–3176, IEEE, 2011.

Y

[Yengui *et al.*, 09] : Yengui A., Neji M., "*Semantic annotation formalism of video-conferencing documents*", International Conference IADIS CELDA 2009 Rome, Italy, 20-22 novembre 2009, ISBN: 978-972-8924-95-9 © 2009 IADIS, Acte CELDA Pages 410-414.

[Yengui *et al.*, 10] : Yengui A., Neji M., "*OSVIRA: Ontology-based System for Semantic Visio-conference Information Retrieval and Annotation*", 14th IBIMA June, 2010, Istanbul, ISBN: 978-0-9821489-3-8.

[Yengui *et al.*, 11] : Yengui A., Neji M., "*The Use of Conceptual Graphs to Annotate and to Search Semantically the Pedagogic Videoconferencing*", KCESS, April, 2011, Kuwait.

[Yengui *et al.*, 14] : Yengui A., Neji M., "*Information Research in indexed medical video conferences databases*", IJETTCS, July – August, 2014, Volume 3, Issue 4, ISSN 2278-6856, pages n°226-232.

Z

[Zargayouna *et al.*, 04] : Zargayouna H., Slotti S., "*Mesure de similarité dans une ontology pour l'indexation sémantique de documents XML*", The 5th International Conference on Internet Computing (IC 2004), June 21th - 24th, 2004, Monte Carlo Resort, Las Vegas, Nevada, USA.

[Zhang *et al.*, 06] : Zhang B., Dou W., Chen L., "*Combining short and long term audio features for tv sports highlight detection*", In Advance In Information Retrieval, 28th European Conference on IR Research, ECIR 2006, Lecture Notes in Computer Science, pages 472-475, London, UK, 2006, Springer.

[Zemirli, 08] : Zemirli N., "*Modèle d'accès personnalisé à l'information basé sur les Diagrammes d'Influence intégrant un profil utilisateur évolutif*", Thèse de l'Université Paul Sabatier de Toulouse III, soutenue le 12 Juin 2008.

[Zhou *et al.*, 07] : Zhou X., Shen H.T., "*A new similarity measure for near duplicate video clip detection*", In Advances in Data and Web Management, Joint 9th Asia-Pacific Web Conference, APWeb 2007, and 8th International Conference, on Web-Age Information Management, WAIM 2007, Volume 4505 of Lecture Notes in Computer Science, pages 176-187, Huang Shan, China, 2007. Springer.

[Zghal, 10] : Zghal S., "*Contributions à l'alignement d'ontologies OWL par agrégation de similarités*", Thèse de l'Université de Tunis El Manar et l'université D'Artois Soutenue le 21 décembre 2010.

[Zwidi *et al.*, 13] : Zwidi A., Yengui A., Neji M., "*Research system of semantic information in medical videoconference based on conceptual graphs and domain ontologies*", IJMIT November, 2013, Vol. 7, No. 2, ISSN 2278-5612, Page 969.

Glossaire

Terme	Signification
ADM	<u>A</u> ide au <u>D</u> iagnostic <u>M</u> édical
Ao	<u>O</u> bjets <u>A</u> udio
API	<u>A</u> pplication <u>P</u> rogram <u>I</u> nterface
Asc	<u>C</u> oncept <u>S</u> émantique <u>A</u> udio
AV	<u>A</u> udio <u>v</u> isuel
BD	<u>B</u> ase de <u>D</u> onnées
CA	<u>C</u> lassification <u>A</u> utomatique
CBIR	<u>C</u> ontent <u>B</u> ased <u>I</u> nformation <u>R</u> etrieval
Ccat	<u>C</u> olor <u>cat</u> egory
CF	<u>F</u> réquence d'un <u>C</u> oncept
CID	<u>C</u> ontext based <u>I</u> nference <u>D</u> iagram
CMYK	<u>C</u> yan, <u>M</u> agenta, <u>Y</u> ellow, <u>K</u> ey
COMP	Relation conceptuelle de <u>com</u> position
CVA	<u>C</u> ognitive <u>V</u> isual <u>A</u> ttention
DAML	<u>D</u> ARPA <u>A</u> gent <u>M</u> arkup <u>L</u> anguage
DAML+OIL	Un langage ontologique pour le Web sémantique. Il est le successeur du langage DAML et OIL en combinant les caractéristiques des deux langages
DAV	<u>D</u> ocuments <u>A</u> udio <u>v</u> isuels
DWT	<u>D</u> iscrete <u>W</u> avelet <u>T</u> ransform
EMIR ²	An <u>E</u> xtended <u>M</u> odel for <u>I</u> mage <u>R</u> epresentation and <u>R</u> etrieval
G	<u>G</u> raphe
GCs	<u>G</u> raphes <u>C</u> onceptuels
GO	<u>G</u> ene <u>O</u> ntology
GPS	<u>G</u> lobal <u>P</u> ositioning <u>S</u> ystem
HMM	<u>H</u> idden <u>M</u> arkov <u>M</u> odels
HP	<u>H</u> ewlett- <u>P</u> ackard
IA	<u>I</u> ntelligence <u>A</u> rtificielle
IBM	<u>I</u> nternational <u>B</u> usiness <u>M</u> achines
Ids	<u>D</u> escripteurs <u>I</u> mages

IHM	<u>I</u> nteraction <u>H</u> ome- <u>M</u> achine
IMA	<u>I</u> nformation <u>M</u> utuelle <u>A</u> daptée
Io	<u>O</u> bjet <u>I</u> mage
IPN	<u>I</u> ndicateur <u>P</u> référé ou <u>N</u> on
IUC	<u>I</u> dentificateur <u>U</u> nique du <u>C</u> oncept
IUTS	<u>I</u> dentificateur <u>U</u> nique du <u>T</u> ype <u>S</u> émantique
JVA	<u>J</u> ava <u>V</u> irtual <u>M</u> achine
K-SPACE	<u>K</u> nowledge <u>S</u> pace of Semantic Interface for Automatic Indexation and Retrieval of Multimedia Content
LVCSR	<u>L</u> arge <u>V</u> ocabulary <u>C</u> ontinuous <u>S</u> peech <u>R</u> ecognition
MAP	<u>M</u> ean <u>A</u> verage <u>P</u> recision
MDV	<u>M</u> edical <u>V</u> isio-conference <u>D</u> ocuments
MeSH	<u>M</u> edical <u>S</u> ubject <u>H</u> eading
MFCC	<u>M</u> el- <u>F</u> requency <u>C</u> epstral <u>C</u> oefficient
MHM	Maladies <u>H</u> éréditaires du <u>M</u> étabolisme
MMM ²	<u>M</u> obile <u>M</u> edia <u>M</u> etadata for <u>M</u> edia sharing
MPEG-7	<u>M</u> oving <u>P</u> icture <u>E</u> xperts <u>G</u> roup
MUSCL	<u>M</u> ultimedia <u>U</u> nderstanding through <u>S</u> emantics, <u>C</u> omputation and <u>L</u> earning
NCG (TNG)	<u>N</u> ested <u>C</u> onceptual <u>G</u> raphs
OGMS	The <u>O</u> ntology for <u>G</u> eneral <u>M</u> edical <u>S</u> cience
OIL	<u>O</u> ntology <u>I</u> nterchange <u>L</u> anguage
OntoEdit	<u>O</u> ntology <u>E</u> ditor
ONTOMÉNÉLAS	<u>O</u> ntology which has been developed during the <u>M</u> énélas project in the cardiac surgery domain
OS	<u>O</u> ntology <u>S</u> ource
OSSVIRI	<u>O</u> ntology-based <u>S</u> ystem for <u>S</u> emantic <u>V</u> isio-conference <u>I</u> nformation <u>R</u> etrieval and <u>I</u> ndexation
OWL	<u>W</u> eb <u>O</u> ntology <u>L</u> anguage
PDA	<u>P</u> ersonal <u>D</u> igital <u>A</u> ssistant
PU	<u>P</u> rofil <u>U</u> tilisateur
RAP	<u>R</u> econnaissance <u>A</u> utomatique de la <u>P</u> arole
RDF	<u>R</u> esource <u>D</u> escription <u>F</u> ramework
RGB / RVB	<u>R</u> ed, <u>G</u> reen, <u>B</u> lue
RE	<u>R</u> essource <u>E</u> xterne
RI	<u>R</u> echerche d' <u>I</u> nformations
RS	<u>R</u> elation <u>S</u> émantique

RSV	<u>R</u> etrieval <u>S</u> tatus <u>V</u> alue
SG	<u>G</u> raphe Conceptuel <u>S</u> imple
SGBDR	<u>S</u> ystème de <u>G</u> estion de <u>B</u> ase de <u>D</u> onnées <u>R</u> elationnelles
SID	<u>S</u> imple <u>I</u> nterface <u>D</u> igram
SLR	<u>S</u> egmentation et <u>R</u> egroupement en <u>L</u> ocuteurs
SMI de Stanford	<u>S</u> tanford's <u>M</u> edical <u>I</u> nformatics <u>S</u> ection
SpC	treillis de relations <u>S</u> patiales <u>C</u> onceptuelles
SQL	<u>S</u> tructured <u>Q</u> uery <u>L</u> anguage. Language de requête pour des bases de Données relationnelles
SRI	<u>S</u> ystème de <u>R</u> echerche d' <u>I</u> nformation
STFT	<u>S</u> hort <u>T</u> erm <u>F</u> ourier <u>T</u> ransform
SVD	<u>D</u> écomposition par <u>V</u> aleur <u>S</u> ingulière
SVM	<u>M</u> achine à <u>V</u> ecteurs de <u>S</u> upport
TA	<u>T</u> erme <u>A</u> ssocié
TAC	<u>T</u> erms <u>A</u> ssociés au <u>C</u> oncept
TAL	<u>T</u> raitement <u>A</u> utomatique des <u>L</u> angues
TALN	<u>T</u> raitement <u>A</u> utomatique du <u>L</u> angage <u>N</u> aturel
TeR	<u>R</u> elations <u>T</u> emporelles
TextureLEP	<u>T</u> exture <u>L</u> ocal- <u>E</u> dge <u>P</u> attern
TG	<u>T</u> erme <u>G</u> énérique
TIC	<u>T</u> echnologie de l' <u>I</u> nformation et de <u>C</u> ommunication
TS	<u>T</u> erme <u>S</u> pécifique
TT	<u>T</u> hème
UAV	<u>U</u> nité <u>A</u> udio <u>V</u> isuelle
UMLS	<u>U</u> nified <u>M</u> edical <u>L</u> anguage <u>S</u> ystem
URL	<u>U</u> ser <u>R</u> equirement <u>L</u> anguages
VS	<u>V</u> ideo <u>S</u> egment
VSC	<u>V</u> isual <u>S</u> emantic <u>C</u> oncept
XML	e <u>X</u> tensible <u>M</u> arkup <u>L</u> anguage



Annexes

XML Visioconférence

```
<?xml version="1.0" encoding="utf-8"?>
<Vidéo1>
<CompleteName>
<titre>Vidéo Ina Vidéo chirurgie </titre>
<Format>AVI</Format>
<Emplacement>C:\Users\hp\Desktop\BASE
DAV\visioconfdecompress201.avi</Emplacement>
</CompleteName>
<Contenu>
<audio>
<ligne>It happened for the first time in the world.</ligne>
<ligne> The Professor Carpentier's team in Paris managed an
open heart surgery by video surgery, i.e by means of a small
camera inserted into the heart after a small incision.
</ligne>
<ligne> In the heart, between the atrium and the left
ventricle, the surgeon's act is of absolute control.</ligne>
<ligne> A microvideo camera has guided the intervention.
</ligne>
<ligne> the image is enlarged to facilitate the work, a
surgical stitches to repair a damaged heart </ligne>
<ligne> he patient was thirty years old and the operation time
is two hours and a half. </ligne>
<ligne> It is like a spirologue, a spirologue surgeon.</ligne>
<ligne> we go down into a chest cavity, and then into another
heart cavity. </ligne>
<ligne> There, we find enlarged, by 4 to 5 times, the
structures that we know and some can never seen because they
cannot be seen with normal vision.</ligne>
<ligne> So it is very moving.</ligne>
<ligne> Without the help of video surgery, it would have
required to open the patient's trunk as shown in this
diagram on the left of the screen. </ligne>
<ligne> For this procedure the incision is only four
centimeters.</ligne>
<ligne> A first advantage: The operation shock is much less
important. There is less blood loss and less risk of
infection. </ligne>
</audio>
</Contenu>
</Vidéo1>
```


Liste des mots vides

```

<?xml version="1.0"
encoding="utf-8"?>
<MotVide>
<mot>a</mot>
<mot>about</mot>
<mot>above</mot>
<mot>across</mot>
<mot>actually</mot>
<mot>after</mot>
<mot>afterwards</mot>
<mot>again</mot>
<mot>against</mot>
<mot>ago</mot>
<mot>all</mot>
<mot>almost</mot>
<mot>alone</mot>
<mot>along</mot>
<mot>already</mot>
<mot>also</mot>
<mot>although</mot>
<mot>always</mot>
<mot>am</mot>
<mot>among</mot>
<mot>amongst</mot>
<mot>amongst</mot>
<mot>amountv</mot>
<mot>an</mot>
<mot>and</mot>
<mot>another</mot>
<mot>any</mot>
<mot>anyhow</mot>
<mot>anyone</mot>
<mot>anything</mot>
<mot>anyway</mot>
<mot>anywhere</mot>
<mot>are</mot>
<mot>around</mot>
<mot>as</mot>
<mot>at</mot>
<mot>back</mot>
<mot>be</mot>
<mot>became</mot>
<mot>because</mot>
<mot>become</mot>
<mot>becomes</mot>
<mot>becoming</mot>
<mot>been</mot>
<mot>before</mot>
<mot>beforehand</mot>
<mot>behind</mot>
<mot>being</mot>
<mot>below</mot>
<mot>beside</mot>
<mot>besides</mot>
<mot>between</mot>
<mot>beyond</mot>
<mot>bill</mot>
<mot>both</mot>
<mot>bottom</mot>
<mot>but</mot>
<mot>by</mot>
<mot>call</mot>
<mot>can</mot>
<mot>cannot</mot>
<mot>cant</mot>
<mot>co</mot>
<mot>computer</mot>
<mot>con</mot>
<mot>could</mot>
<mot>couldnt</mot>
<mot>cry</mot>
<mot>currently</mot>
<mot>de</mot>
<mot>describe</mot>
<mot>detail</mot>
<mot>do</mot>
<mot>done</mot>
<mot>down</mot>
<mot>due</mot>
<mot>during</mot>
<mot>each</mot>
<mot>eg</mot>
<mot>eight</mot>
<mot>either</mot>
<mot>eleven</mot>
<mot>else</mot>
<mot>elsewhere</mot>
<mot>empty</mot>
<mot>enough</mot>
<mot>etc</mot>
<mot>even</mot>
<mot>evening</mot>
<mot>ever</mot>
<mot>every</mot>
<mot>everyone</mot>
<mot>everything</mot>
<mot>everywhere</mot>
<mot>except</mot>
<mot>few</mot>
<mot>fifteen</mot>
<mot>fifty</mot>
<mot>fill</mot>
<mot>find</mot>
<mot>fire</mot>
<mot>first</mot>
<mot>five</mot>

```

<mot>for</mot>	<mot>mill</mot>
<mot>for this reason</mot>	<mot>mine</mot>
<mot>former</mot>	<mot>more</mot>
<mot>formerly</mot>	<mot>moreover</mot>
<mot>forty</mot>	<mot>morning</mot>
<mot>found</mot>	<mot>most</mot>
<mot>four</mot>	<mot>mostly</mot>
<mot>from</mot>	<mot>move</mot>
<mot>front</mot>	<mot>much</mot>
<mot>full</mot>	<mot>must</mot>
<mot>further</mot>	<mot>my</mot>
<mot>get</mot>	<mot>myself</mot>
<mot>give</mot>	<mot>name</mot>
<mot>go</mot>	<mot>namely</mot>
<mot>had</mot>	<mot>neither</mot>
<mot>has</mot>	<mot>never</mot>
<mot>hasnt</mot>	<mot>nevertheless</mot>
<mot>have</mot>	<mot>next</mot>
<mot>hence</mot>	<mot>nine</mot>
<mot>her</mot>	<mot>no</mot>
<mot>here</mot>	<mot>nobody</mot>
<mot>hereafter</mot>	<mot>none</mot>
<mot>hereby</mot>	<mot>noone</mot>
<mot>herein</mot>	<mot>nor</mot>
<mot>hereupon</mot>	<mot>normally</mot>
<mot>hers</mot>	<mot>not</mot>
<mot>herself</mot>	<mot>nothing</mot>
<mot>himself</mot>	<mot>now</mot>
<mot>his</mot>	<mot>nowhere</mot>
<mot>how</mot>	<mot>of</mot>
<mot>however</mot>	<mot>off</mot>
<mot>hundred</mot>	<mot>often</mot>
<mot>i</mot>	<mot>on</mot>
<mot>ie</mot>	<mot>once</mot>
<mot>if</mot>	<mot>one</mot>
<mot>in</mot>	<mot>only</mot>
<mot>inc</mot>	<mot>onto</mot>
<mot>indeed</mot>	<mot>or</mot>
<mot>interest</mot>	<mot>other</mot>
<mot>into</mot>	<mot>others</mot>
<mot>is</mot>	<mot>otherwise</mot>
<mot>its</mot>	<mot>our</mot>
<mot>itself</mot>	<mot>ours</mot>
<mot>keep</mot>	<mot>ourselves</mot>
<mot>last</mot>	<mot>out</mot>
<mot>latter</mot>	<mot>over</mot>
<mot>latterly</mot>	<mot>own</mot>
<mot>least</mot>	<mot>part</mot>
<mot>less</mot>	<mot>per</mot>
<mot>ltd</mot>	<mot>perhaps</mot>
<mot>made</mot>	<mot>please</mot>
<mot>many</mot>	<mot>put</mot>
<mot>may</mot>	<mot>rather</mot>
<mot>me</mot>	<mot>re</mot>
<mot>meanwhile</mot>	<mot>same</mot>
<mot>might</mot>	<mot>see</mot>

<mot>seem</mot>	<mot>toward</mot>
<mot>seemed</mot>	<mot>towards</mot>
<mot>seeming</mot>	<mot>twelve</mot>
<mot>seems</mot>	<mot>twenty</mot>
<mot>serious</mot>	<mot>two</mot>
<mot>several</mot>	<mot>un</mot>
<mot>should</mot>	<mot>under</mot>
<mot>show</mot>	<mot>until</mot>
<mot>side</mot>	<mot>up</mot>
<mot>since</mot>	<mot>upon</mot>
<mot>sincere</mot>	<mot>us</mot>
<mot>six</mot>	<mot>very</mot>
<mot>sixty</mot>	<mot>via</mot>
<mot>so</mot>	<mot>we</mot>
<mot>so that</mot>	<mot>was</mot>
<mot>some</mot>	<mot>well</mot>
<mot>somehow</mot>	<mot>were</mot>
<mot>someone</mot>	<mot>what</mot>
<mot>something</mot>	<mot>whatever</mot>
<mot>sometime</mot>	<mot>when</mot>
<mot>sometimes</mot>	<mot>whence</mot>
<mot>somewhere</mot>	<mot>whenever</mot>
<mot>soon</mot>	<mot>where</mot>
<mot>still</mot>	<mot>whereafter</mot>
<mot>such</mot>	<mot>whereas</mot>
<mot>take</mot>	<mot>whereby</mot>
<mot>ten</mot>	<mot>wherein</mot>
<mot>than</mot>	<mot>whereupon</mot>
<mot>the</mot>	<mot>wherever</mot>
<mot>their</mot>	<mot>whether</mot>
<mot>them</mot>	<mot>while</mot>
<mot>themselves</mot>	<mot>whither</mot>
<mot>then</mot>	<mot>whoever</mot>
<mot>thence</mot>	<mot>whole</mot>
<mot>there</mot>	<mot>whom</mot>
<mot>thereafter</mot>	<mot>whose</mot>
<mot>thereby</mot>	<mot>why</mot>
<mot>therefore</mot>	<mot>will</mot>
<mot>therein</mot>	<mot>with</mot>
<mot>thereupon</mot>	<mot>within</mot>
<mot>these</mot>	<mot>without</mot>
<mot>thick</mot>	<mot>would</mot>
<mot>thin</mot>	<mot>yet</mot>
<mot>third</mot>	<mot>you</mot>
<mot>this</mot>	<mot>your</mot>
<mot>those</mot>	<mot>yours</mot>
<mot>though</mot>	<mot>yourself</mot>
<mot>three</mot>	<mot>yourselves</mot>
<mot>through</mot>	<mot>a</mot>
<mot>throughout</mot>	<mot>about</mot>
<mot>thru</mot>	<mot>an</mot>
<mot>thus</mot>	<mot>are</mot>
<mot>to</mot>	<mot>as</mot>
<mot>together</mot>	<mot>at</mot>
<mot>too</mot>	<mot>be</mot>
<mot>top</mot>	<mot>by</mot>

<mot>com</mot>
<mot>de</mot>
<mot>en</mot>
<mot>for</mot>
<mot>from</mot>
<mot>how</mot>
<mot>in</mot>
<mot>is</mot>
<mot>it</mot>
<mot>la</mot>
<mot>is</mot>
<mot>of</mot>
<mot>on</mot>
<mot>or</mot>
<mot>the</mot>
<mot>this</mot>
<mot>to</mot>
<mot>was</mot>
<mot>what</mot>
<mot>when</mot>
<mot>where</mot>
<mot>who</mot>
<mot>will</mot>
<mot>with</mot>
<mot>und</mot>
<mot>the</mot>
<mot>www</mot>
</MotVide>

Liste des Termes Simples

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<racine>
```

```
<term>six</term>
```

```
<term>heart</term>
```

```
<term>surgeries</term>
```

```
<term>last</term>
```

```
<term>week</term>
```

```
<term>robot</term>
```

```
<term>fantastic</term>
```

```
<term>technical</term>
```

```
<term>progress</term>
```

```
<term>recent</term>
```

```
<term>years</term>
```

```
<term>innovations</term>
```

```
<term>Tele</term>
```

```
<term>surgery</term>
```

```
<term>mini</term>
```

```
<term>incision</term>
```

```
<term>millimetres</term>
```

```
<term>French</term>
```

```
<term>technology</term>
```

```
<term>surgeons</term>
```

```
<term>professional</term>
```

```
<term>Institute</term>
```

```
<term>Research</term>
```

```
<term>Cancer</term>
```

```
<term>Strasbourg</term>
```

```
<term>surgical</term>
```

```
<term>act</term>
```

```
<term>intra</term>
```

```
<term>abdominal</term>
```

```
<term>colon</term>
```

```
<term>tens</term>
```

```
<term>London</term>
```

```
<term>Munich</term>
```

```
<term>delicate</term>
```

```
<term>interventions</term>
```

```
<term>live</term>
```

```
<term>videoconference</term>
```

```
<term>160</term>
```

```
<term>specialists</term>
```

```
<term>laparoscopic</term>
```

```
<term>conference</term>
```

```
<term>room</term>
```

```
<term>eye</term>
```

```
<term>camera</term>
```

```
<term>patient</term>
```

```
<term>digestive</term>
```

```
<term>system</term>
```

```
<term>precision</term>
```

```
<term>obstacle</term>
```

```
<term>codification</term>
```

```
<term>reason</term>
```

<term>meetings</term>
<term>transfer</term>
<term>messages</term>
<term>reproducible</term>
<term>acts</term>
<term>hands</term>
<term>3000</term>
<term>Professor</term>
<term>Leroy</term>
<term>Joiil</term>
<term>performs</term>
<term>youth</term>
<term>Japan</term>
<term>he</term>
<term>France</term>
<term>new</term>
<term>techniques</term>
<term>teaches</term>
<term>theory</term>
<term>course</term>
<term>practice</term>
<term>tables</term>
<term>ports</term>
<term>slaughter</term>
<term>house</term>
<term>cancer</term>
<term>experts</term>
<term>immediate</term>
<term>response</term>
<term>video</term>
<term>computer</term>
<term>third</term>
<term>millennium</term>
<term>medicine</term>
<term>Bourgeois</term>
<term>prepared</term>
<term>decade</term>
<term>conservative</term>
<term>future</term>
<term>advantage</term>
<term>researchers</term>
<term>world</term>
<term>Ina</term>
<term>fr</term>
<term>University</term>
<term>Hospital</term>
<term>European</term>
<term>operating</term>
<term>Doctor</term>
<term>Miyake</term>
<term>York</term>
<term>Jeffrey</term>
<term>WWilson</term>
<term>Jacques</term>
<term>Marescaux</term>

<term>Pdt</term>
<term>IRCAD</term>
 <term>achieved</term >
 <term>trained</term>
 <Term>codified</Term>
 <Term>practice</Term>
 <Term>intended</Term>
 <Term>treated</Term>
 <Term>prepared</Term>
 <Term>generalized</Term>
 <Term>assist</Term>
 <Term>systematized</Term>
</racine>

Liste des Termes Composés

```
<?xml version="1.0" encoding="UTF-8"?>
<racine>
<Terme_composi@>six heart surgeries</Terme_composi@>
<Terme_composi@>occurred last week</Terme_composi@>
<Terme_composi@>fantastic technical progress</Terme_composi@>
<Terme_composi@>recent years</Terme_composi@>
<Terme_composi@>Tele surgery</Terme_composi@>
<Terme_composi@>mini incision</Terme_composi@>
<Terme_composi@>French technology</Terme_composi@>
<Terme_composi@>Institute Research</Terme_composi@>
<Terme_composi@>Cancer Strasbourg</Terme_composi@>
<Terme_composi@>surgical act</Terme_composi@>
<Terme_composi@>intra abdominal colon incision</Terme_composi@>
<Terme_composi@>tens millimetres</Terme_composi@>
<Terme_composi@>160@i specialists</Terme_composi@>
<Terme_composi@>laparoscopic surgery</Terme_composi@>
<Terme_composi@>conference room</Terme_composi@>
<Terme_composi@>includes 160 specialists</Terme_composi@>
<Terme_composi@>surgeons eye</Terme_composi@>
<Terme_composi@>mini camera</Terme_composi@>
<Terme_composi@>patient digestive system</Terme_composi@>
<Terme_composi@>surgical act</Terme_composi@>
<Terme_composi@>surgical act</Terme_composi@>
<Terme_composi@>years 3000 surgeons</Terme_composi@>
<Terme_composi@>French technology</Terme_composi@>
<Terme_composi@>Professor Leroy Joil</Terme_composi@>
<Terme_composi@>performs laparoscopic surgery</Terme_composi@>
<Terme_composi@>new techniques</Terme_composi@>
<Terme_composi@>Institute Tele surgery Strasbourg</Terme_composi@>
<Terme_composi@>surgery tables</Terme_composi@>
<Terme_composi@>slaughter house</Terme_composi@>
<Terme_composi@>colon cancer</Terme_composi@>
<Terme_composi@>immediate response</Terme_composi@>
<Terme_composi@>video surgery</Terme_composi@>
<Terme_composi@>computer surgery</Terme_composi@>
<Terme_composi@>third millennium medicine</Terme_composi@>
<Terme_composi@>institute Strasbourg</Terme_composi@>
<Terme_composi@>laparoscopic surgery</Terme_composi@>
<Terme_composi@>conservative surgeons</Terme_composi@>
<Terme_composi@>French technology</Terme_composi@>
<Terme_composi@>Ina fr</Terme_composi@>
<Terme_composi@>University Hospital Strasbourg</Terme_composi@>
<Terme_composi@>European Institute Tele surgery</Terme_composi@>
<Terme_composi@>operating room</Terme_composi@>
<Terme_composi@>includes surgeons</Terme_composi@>
<Terme_composi@>Professor Leroy Joil</Terme_composi@>
<Terme_composi@>Doctor Miyake</Terme_composi@>
<Terme_composi@>New York</Terme_composi@>
<Terme_composi@>Doctor Jeffrey WWilson</Terme_composi@>
<Terme_composi@>Professor Jacques Marescaux</Terme_composi@>
<Terme_composi@>Pdt IRCAD European institute</Terme_composi@>
<Terme_composi@>Tele surgery</Terme_composi@>
</racine>
```


Liste des Concepts

```
<?xml version="1.0" encoding="UTF-8"?>
<racine>
<exemple>
<concept>Tele surgery</concept>
<concept>surgical act</concept>
<concept>160 specialists</concept>
<concept>laparoscopic surgery</concept>
<concept>Professor Leroy Joiil</concept>
<concept>European Institute Tele surgery</concept>
<concept>surgery</concept>
<concept>surgeons</concept>
<concept>act</concept>
<concept>Doctor Jeffrey WWilson</concept>
<concept>Doctor Miyake</concept>
<concept>Institute Tele surgery Strasbourg</concept>
<concept>Professor Jacques Marescaux</concept>
<concept>computer surgery</concept>
<concept>last week</concept>
<concept>intra abdominal colon incision</concept>
<concept>mini camera</concept>
<concept>mini incision</concept>
<concept>operating room</concept>
<concept>patient digestive system</concept>
<concept>six heart surgeries</concept>
<concept>surgeons eye</concept>
<concept>third millennium medicine</concept>
<concept>video surgery</concept>
<concept>future</concept>
</exemple>
</racine>
```

Liste des relations sémantiques

```
<?xml version="1.0" encoding="UTF-8"?>
<racine>
  <Semantic_R>occur</Semantic_R>
  <Semantic_R>allow</Semantic_R>
  <Semantic_R>consist_of</Semantic_R>
  <Semantic_R>include</Semantic_R>
  <Semantic_R>introduce</Semantic_R>
  <Semantic_R>use</Semantic_R>
  <Semantic_R>come</Semantic_R>
  <Semantic_R>say</Semantic_R>
  <Semantic_R>happen</Semantic_R>
  <Semantic_R>optimize</Semantic_R>
  <Semantic_R>mean</Semantic_R>
  <Semantic_R>highlight</Semantic_R>
  <Semantic_R>teach</Semantic_R>
  <Semantic_R>hold</Semantic_R>
  <Semantic_R>learn</Semantic_R>
  <Semantic_R>answer</Semantic_R>
  <Semantic_R>fear</Semantic_R>
  <Semantic_R>happen</Semantic_R>
</racine>
```

Exemple de requêtes traitées

- R15 les maladies du cœur causent le rythme raide du cœur
- R21 le traitement du rachitisme carenciel
- R24 le traitement proposé dans le cas d'une hémorragie digestive
- R25 traiter les complications de l'angioplastie à la phase aiguë de l'infarctus du myocarde
- R26 traiter le schwannome vestibulaire
- R27 traitement des métastases cérébrales
- R31 le traitement de l'hypernatrémie
- R37 traiter un cancer du rein
- R50 technique peut permettre de détecter une tumeur des tissus