

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

وزارة التعليم العالي و البحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITY BADJI MOKHTAR-ANNABA



جامعة باجي مختار- عنابة

Faculté des Sciences de l'Ingénierat
Département d'Informatique

Année : 2013/2014

THESE

Présentée en vue de l'obtention du diplôme de

Doctorat 3^{ième} Cycle LMD

Intitulée :

**Préparation et organisation d'un domaine d'apprentissage à partir
du Web pour un enseignement adapté au profil de l'apprenant**

Option
Informatique

Par
Mohammed CHAOUI

Devant le Jury

Pr Nadir FARAH	Professeur à l'Université d'Annaba	Président
Pr Mohamed Tayeb LASKRI	Professeur à l'Université d'Annaba	Directeur de Thèse
Pr Nacer eddine ZAROUB	Professeur à l'Université de Constantine	Examineur
Pr Azeddine BILAMI	Professeur à l'Université de Batna	Examineur
Pr Hayet Farida MEROUANI	Professeur à l'Université d'Annaba	Examineur

A mon encadreur.
A mes Parents.
A mon Frère.
A ma Sœur.
A ma Fiancée.
A mes Collègues.
A ma Famille.
A mes Amis.

Remerciements

Je remercie en premier lieu mon grand DIEU qui m'a donné à la fois le courage, la volonté, et la patience afin d'élaborer cette thèse de recherche scientifique.

- *Je tiens à remercier, mon directeur de thèse Monsieur Mohamed Tayeb Laskri, Professeur à l'université Badji Mokhtar à Annaba pour l'encadrement de mon travail et pour son encouragement, ainsi que son soutien tout au long de la thèse. Je le remercie pour tout son aide. Son enthousiasme et sa patience ont beaucoup facilité et agrémenté mon travail. Il a été toujours disponible pour répondre aux questions que je lui posais. Ses remarques m'ont permis de faire progresser ce travail.*
- *Je tiens également à remercier les membres du jury qui m'ont fait l'honneur de bien vouloir évaluer mon travail, et plus précisément :*
 - *Monsieur Nadir FARAH, Professeur à l'Université d'Annaba, pour l'honneur qu'il m'a fait, en acceptant la présidence de ce jury.*
 - *Monsieur Nacer eddine ZAROUR, Professeur à l'Université de Constantine, Monsieur Azeddine BILAMI, Professeur à l'Université de Batna et Madame Hayet Farida MEROUANI, Professeur à l'Université d'Annaba pour avoir accepté de juger le présent document.*
- *Je tiens à remercier mes parents pour leurs encouragements et soutiens pendant toute la période de la thèse en leurs souhaitant une belle vie.*
- *Je tiens à remercier mon frère, ma sœur, ma fiancée et toute la famille.*
- *Enfin, je tiens à remercier mes amis et mes collègues, en leurs espérant une bonne continuation et une bonne chance.*

ملخص

يرتبط مجال التعليم بصفة كبيرة مع تكنولوجيات الإعلام والاتصال التي تكيفت مع متطلبات تعليمية. هذه التكنولوجيات أتاحت موضوع دراسة مشتركة في عدة مجالات منها الإعلام الآلي، العلوم المعرفية والعلوم التعليمية. لحد الساعة، نلاحظ وجود هدف مهم جدا يتطور من يوم إلى آخر يتعلق بتكنولوجيات الشبكات والذي يحتل فيه الويب مكان رئيسي.

يعتبر الويب أو شبكة الانترنت مجالا للإنتاج، التسيير و الاتصال للمحتوى التعليمي في تطبيقاتية التعلم الإلكتروني. هذا المجال يتطور بسرعة كبيرة و يتأثر بعوامل ترتبط أساسا بتشغيله و معناه. نفس المجال يفرض مصاعب للأساتذة خاصة في استخلاص المعلومات الملائمة. إذن الويب يسمح بالتصفح المستقل من أجل هدف تعليمي من خلال الاتصال مع الكميات الهائلة للموارد المتواجدة في الويب. بالاعتماد على وسائل الويب، قد يتعب الأساتذة في عملية البحث و تصفية معلومات الويب خاصة إذا واجهتهم صيغ مختلفة لموارد الويب مع بنيات معقدة. تكيف موارد الويب المستخلصة يجب أن يؤمن من طرف الأساتذة لإعطاء الدقة و رضا المتكويين و كذا إعطاء أحسن نوعية لمحتوى قاعدة التعلم الإلكتروني.

هذه الدراسة تقترح طريقة دمج لموارد الويب لهذا الإشكال و كذا تنظيم هذه الموارد في شكل مجال تعليمي مكيف داخل قاعدة التعلم الإلكتروني. الهدف مركز في مرحلة أولى على وسائل البحث و طرق التصفية من أجل استخلاص موارد الويب الملائمة مع تنظيم هذه الأخيرة من أجل إنشاء الدروس. تتركز المرحلة الثانية على تكيف موارد الويب. هذه الطريقة تستكشف عملية دمج جديدة في ميدان إنشاء الدروس و كذا تكيفها حسب الملفات الشخصية للمتعلمين. هذه الطريقة لا تتطلب وقت كثير أو مجهود كبير اللذين قد يبذلها الأستاذ من أجل إنشاء الدروس. بالإضافة، هذه الطريقة تجد بسهولة المنهج الأمثل من أجل إيجاد موارد الويب المطلوبة. عملية تحديث الدروس تتم مباشرة عن طريق الويب و خاصة إعادة استغلال موارد الويب المستخلصة من قبل. تقييم هذا العمل أعطى أداء جيد مقارنة مع مختلف طرق إنشاء الدروس.

الكلمات الدلالية

قاعدة التعلم الإلكتروني، موارد الويب، أنطولوجيا المجال، أنطولوجيا تربوية، مجال تعليمي مكيف، الدمج و الملفات الشخصية للمتعلمين.

Résumé

Le domaine d'enseignement a toujours été en relation étroite avec les technologies de l'information et de la communication (TIC). Celles-ci ont été adaptées à des fins didactiques, et leurs capacités ont fait le sujet d'étude de diverses disciplines, à savoir l'informatique, les sciences cognitives et les sciences d'enseignement. À l'heure actuelle, nous constatons qu'un rôle de plus en plus important est accordé aux technologies de réseau dont le Web trouve une place centrale.

Dans une application d'enseignement en ligne, le Web présente un espace pour la production, la gestion et la communication du contenu éducatif. Cet espace évolue rapidement ; il est régi par des facteurs liés à son fonctionnement et à sa signification. Ce même espace pose des difficultés aux enseignants, surtout pour l'extraction des informations les plus pertinentes. Le Web donc, permet l'auto-navigation pour un objectif éducatif à travers l'interaction avec des grandes quantités de ressources Web. Tout en profitant de la souplesse des outils Web, les auteurs peuvent souffrir dans la recherche et le filtrage des ressources Web, quand ils sont devant des formats différents des ressources Web avec des structures complexes. Une adaptation des ressources du Web extraites doit être assurée par les auteurs, pour donner la fiabilité, la satisfaction des apprenants et la meilleure qualité du contenu de la plateforme E-Learning.

Cette étude propose une approche de Fusion des ressources Web à ce problème, et organise ces ressources en un domaine d'apprentissage adaptatif dans la plateforme E-Learning. L'accent est concentré d'une part, sur les outils de recherche et les méthodes de filtrage pour extraire les ressources Web éducatives les plus pertinents, et les structurer pour créer les cours. D'autre part, sur l'adaptation des ressources Web extraites. Notre approche explore un nouveau processus de fusion dans la création et dans l'adaptation aux profils des apprenants. Cette approche n'exige pas beaucoup de temps et nombreux efforts qui peuvent être prises et réalisées par des auteurs pour créer les cours. Elle trouve également un chemin direct aux ressources Web qui sont nécessaires, et la mise à jour des cours qui peut être effectuée directement à partir du Web avec la réutilisation des ressources extraites. L'évaluation de ce travail a donné des performances élevées en comparaison avec les différentes méthodes de création des cours.

Mots clés

Plateformes d'enseignement en ligne ou E-Learning, Ressources Web, Ontologie de Domaine, Ontologie Pédagogique, Domaine d'Apprentissage Adaptatif, Fusion, et Profils des Apprenants.

Abstract

The field of education has always been closely connected with information and communication technologies (ICT). These have been adapted for educational purposes, and capacities have been the subject of study in various disciplines, namely computer science, cognitive science and science education. Currently, we perceive that digital and network technologies increase in their importance where the Web plays a central role.

In an application of online learning, Web is used as representative supports for producing, managing and distributing content. This space is evolving rapidly and is governed by factors regarding its function and modes of signification. This same space poses difficulties for teachers, especially for extracting the most relevant information's. The Web allows self-navigated education through interaction with large amounts of Web resources. While enjoying the flexibility of Web tools, authors may suffer from research and filtering Web resources, when they face various resources formats and complex structures. An adaptation of extracted Web resources must be assured by authors, to give reliability, satisfaction of learners and content quality of E-learning platform.

This study has proposed Fusion of Web resources approach to this problem and organized resources to an Adaptive Learning Domain into E-learning platform. The focus was firstly, on searching tools and filtering methods to extract the most relevant educational Web resources and structuring them to create courses. Secondly, on adaptation of extracted Web resources. Our approach explores a new process of fusion in creation and in adaptation to learner's profiles. That approach doesn't need much time and many efforts which can be taken and done by authors to create courses. It also finds direct way to Web resources which are needed, and the update of courses that can be done directly from the Web with the reuse of extracted resources. The evaluation of this work has given high performance in comparison with different methods of course's creation.

Keywords

E-learning Platform, Web resources, Ontology of Domain, Pedagogical Ontology, Adaptive Learning Domain, Fusion, and Learner Profiles.

Table des Matières

Remerciements.....	iii
ملخص.....	iv
Résumé	v
Abstract.....	vi
Table des Matières.....	vii
Table des Illustrations	xiii
Liste des Figures	xiii
Liste des Tableaux	xiv
Liste des Formules	xiv
1. Introduction générale	1
1.1. Contexte et cadre de la recherche	1
1.2. Objectif et approche	2
1.3. Plan du mémoire	5
PARTIE – I : ETAT DE L’ART.....	8
2. La recherche documentaire dans le Web et le filtrage d’information ..	9
2.1. Introduction.....	9
2.2. Notions et concepts de base	10
2.2.1. Recherche d’informations vs Recherche documentaire	10
2.2.2. Filtrage des informations dans le Web.....	10
2.2.3. Annuaire.....	11
2.2.4. Moteur de recherche.....	11
2.2.5. Méta-moteurs	12
2.2.6. Moteur vs annuaire.....	13
2.2.7. L’exploration ou Crawl	13
2.2.8. Action de recherche	13
2.2.9. L’indexation.....	13
2.2.10. Banque des données.....	13
2.2.10.1. Les bases bibliographiques.....	14
2.2.10.2. Les bases textuelles.....	14
2.2.10.3. Les bases de connaissance (factuelles ou numériques).....	14
2.3. Caractéristiques du Web	15
2.3.1. Le Web est énorme.....	15
2.3.2. Le Web vs Bibliothèque.....	15
2.3.3. Le Web invisible	16
2.4. Les outils de recherche	17
2.4.1. Fonctionnement des outils de recherche	17
2.4.2. Types d’outils de recherche	19
2.4.2.1. Les outils humains	19
2.4.2.2. Les outils automatiques (moteurs de recherche)	21

2.4.2.3. Les agents intelligents.....	23
2.5. Distinguer les difficultés.....	24
2.6. Stratégie standard de recherche	26
2.6.1. Les étapes de recherche.....	26
2.6.2. Quel type d'information recherchez-vous ?	26
2.6.3. Définir les mots-clés	27
2.6.4. Formuler la requête.....	28
2.6.5. Analyser les caractéristiques de son sujet de recherche	28
2.6.6. Choisir le bon outil de recherche	29
2.7. L'évaluation des résultats	30
2.8. Conclusion	36
3. Les systèmes éducatifs en ligne.....	37
3.1. Introduction.....	37
3.2. Définition.....	38
3.3. Aperçu historique.....	39
3.4. Les principes du E-Learning.....	40
3.4.1. Premier principe: utilisation des nouvelles technologies.....	40
3.4.2. Deuxième principe : organisation du contenu.....	41
3.4.3. Troisième principe: le rôle de l'enseignant	42
3.4.4. Quatrième principe: la collaboration entre apprenants	43
3.5. Les acteurs et leurs rôles dans le E-Learning	44
3.6. Les outils dans un E-Learning	45
3.6.1. Les outils de gestion des cours (LMS).....	46
3.6.2. Les outils de création des cours (LCMS).....	47
3.6.3. Les outils de communication et de collaboration.....	48
3.7. Conclusion	49
4. E-Learning et le Web Sémantique 'les ontologies'	51
4.1. Introduction.....	51
4.2. Définition.....	52
4.2.1. Le Web Sémantique	52
4.2.2. L'ontologie.....	53
4.3. Objectif du web sémantique	53
4.4. Les langages du Web Sémantique.....	55
4.4.1. RDF	55
4.4.2. OWL.....	56
4.5. L'intérêt de l'ontologie.....	57
4.5.1. Un besoin générique.....	57
4.5.2. Une aide à la communication	58
4.5.3. Une aide à la conception et à l'utilisation des systèmes d'information.....	58
4.6. Composantes de l'ontologie	58
4.6.1. Les concepts	59
4.6.2. Les relations.....	59
4.6.3. Les fonctions.....	59
4.6.4. Les axiomes	59
4.6.5. Les modèles (ou instances)	59
4.7. Dimensions de classification	59
4.7.1. Typologie selon l'objet de conceptualisation.....	59
4.7.2. Typologie selon le niveau de détail de l'ontologie.....	61

4.7.3.	<i>Typologie selon le niveau de complétude</i>	62
4.8.	Les outils de construction d'ontologie	63
4.8.1.	<i>Les outils dépendants de formalisme de représentation</i>	63
4.8.2.	<i>Les outils indépendants de formalisme de représentation</i>	64
4.9.	Méthodologie et cycle de vie d'une ontologie.....	65
4.9.1.	<i>(Uschold et King, 1995)</i>	66
4.10.	Rôle des ontologies pour les applications E-Learning	68
4.11.	Utilité des ontologies pour le E-Learning	68
4.12.	Les ontologies pour les applications E-Learning	69
4.13.	Conclusion	72
5.	Les techniques d'adaptation dans le E-Learning	74
5.1.	Introduction.....	74
5.2.	Définitions	76
5.2.1.	<i>Adaptation</i>	76
5.2.2.	<i>Adaptabilité</i>	76
5.2.3.	<i>Adaptation, Adaptabilité et Adaptativité</i>	76
5.2.4.	<i>Adaptation et Pédagogie</i>	79
5.2.5.	<i>Flexibilité</i>	79
5.2.6.	<i>Prise en compte de l'expérience</i>	79
5.3.	Les systèmes d'enseignement en ligne et l'adaptation.....	79
5.3.1.	<i>Un système adapté</i>	80
5.3.2.	<i>Un système adaptable</i>	80
5.3.3.	<i>Un système adaptatif</i>	80
5.4.	Critères d'adaptation	81
5.4.1.	<i>Adaptation aux caractéristiques de l'utilisateur</i>	81
5.4.1.1.	Connaissances	81
5.4.1.2.	Objectifs (goals ou tasks).....	82
5.4.1.3.	Références (background) et expérience	82
5.4.1.4.	Préférences	83
5.4.1.5.	Choix.....	83
5.4.1.6.	Traits individuels	83
5.4.2.	<i>Adaptation à l'usage</i>	85
5.4.3.	<i>Adaptation à l'environnement</i>	85
5.5.	Méthodes et techniques d'adaptation des systèmes éducatifs	86
5.5.1.	<i>Paramètres d'adaptation</i>	86
5.5.1.1.	Modèle utilisateur:.....	86
5.5.1.2.	Paramètre non-utilisateur ou variable système.....	86
5.5.2.	<i>Modèle d'adaptation</i>	86
5.5.3.	<i>Éléments d'adaptation</i>	87
5.5.4.	<i>Méthodes et techniques</i>	87
5.5.4.1.	Présentation adaptative	87
5.5.4.2.	Support de navigation adaptative	90
5.6.	Conclusion	92
	PARTIE – II : NOTRE PROJET ProALDoF-Web.....	94
6.	Présentation du ProALDoF-Web.....	95
6.1.	Introduction.....	95
6.2.	Motivation.....	96

6.3.	Description du PrOALDoF-Web	97
6.3.1.	<i>Partie Recherche</i>	98
6.3.2.	<i>Partie Fusion</i>	99
6.3.3.	<i>Partie Adaptation</i>	99
6.4.	Architecture de l'approche proposée 'PrOALDoF-Web'	100
6.5.	Description algorithmique du fonctionnement de PrOALDoF-Web	103
6.6.	Conclusion	105
7.	La Recherche et la représentation des connaissances dans PrOALDoF-Web	106
7.1.	Introduction.....	106
7.2.	La Recherche documentaire	106
7.3.	Les Type de documents	107
7.4.	Les sources documentaires	107
7.5.	Etapes de la recherche	108
7.5.1.	<i>La formulation automatique de la requête de recherche</i>	109
7.5.2.	<i>Trouver les informations pertinentes</i>	109
7.5.2.1.	Quel type d'information.....	110
7.5.2.2.	Juger la pertinence des informations	110
7.5.2.3.	L'API Google et la recherche.....	111
7.5.2.4.	Sauvegarde des informations	113
7.5.3.	<i>Analyser le contenu des informations trouvées</i>	114
7.6.	Représentation des connaissances 'Ontologie de domaine'	115
7.6.1.	<i>Représentation est un substitut</i>	115
7.6.2.	<i>Représentation est un ensemble d'engagements ontologiques</i>	115
7.6.3.	<i>Représentation est une théorie fragmentaire du raisonnement intelligent</i> ...	116
7.7.	Description du Domaine d'application.....	116
7.8.	La méthodologie ontologique suivie : Uschold et King's	117
7.8.1.	<i>Identification des objectifs et du contexte</i>	118
7.8.2.	<i>Construction de l'ontologie</i>	119
7.8.2.1.	Première activité : 'capture de l'ontologie'	119
7.8.2.2.	Deuxième activité : 'codage de l'ontologie'	120
7.8.2.3.	Troisième activité : 'intégration d'ontologies existantes'	120
7.8.3.	<i>Evaluation de l'ontologie</i>	120
7.8.4.	<i>Documentation de l'ontologie</i>	121
7.9.	Intégration de l'ontologie de domaine.....	121
7.9.1.	<i>L'API JENA</i>	121
7.9.2.	<i>La hiérarchie du cours</i>	121
7.9.3.	<i>Transformation de l'ontologie en fichier Excel</i>	122
7.10.	Filtrage des informations trouvées.....	124
7.10.1.	<i>Extraction d'informations</i>	124
7.10.2.	<i>Degré de Pertinence 'DP'</i>	126
7.10.3.	<i>Distance Basée sur les Règles Sémantiques 'DBRS'</i>	128
7.10.4.	<i>Précision d'une association 'PA'</i>	130
7.11.	Conclusion	131
8.	La Fusion et l'Adaptation dans PrOALDoF-Web	132
8.1.	Introduction.....	132
8.2.	La Fusion dans PrOALDoF-Web	132

8.3.	Les étapes de la Fusion.....	132
8.3.1.	<i>Créer le modèle Excel</i>	133
8.3.2.	<i>Le processus de la Fusion</i>	136
8.3.3.	<i>L'approche de la Fusion des sous fragments</i>	138
8.3.3.1.	Le niveau d'un Fragment.....	139
8.3.3.2.	Le niveau d'un Sous Fragment.....	139
8.4.	Construction du domaine d'apprentissage 'DA'.....	140
8.5.	L'Adaptation dans PrOALDoF-Web.....	141
8.6.	Un domaine d'apprentissage adaptatif 'DAA'.....	143
8.7.	Conclusion.....	145
9.	Evaluation du PrOALDoF-Web.....	146
9.1.	Introduction.....	146
9.2.	Exemple illustratif.....	147
9.2.1.	<i>Partie 1</i>	148
9.2.1.1.	SF1 par rapport RS1, RS2 et RS3.....	148
9.2.1.2.	SF2 par rapport RS1, RS2 et RS3.....	149
9.2.2.	<i>Partie 2</i>	149
9.2.2.1.	SF1 par rapport RS1, RS2 et RS3 :.....	149
9.2.2.2.	SF2 par rapport RS1, RS2 et RS3.....	150
9.2.3.	<i>Partie 3</i>	150
9.2.3.1.	SF1 par rapport RS1, RS2 et RS3.....	150
9.2.3.2.	SF2 par rapport RS1, RS2 et RS3.....	151
9.2.4.	<i>Partie 4</i>	153
9.2.5.	<i>Partie 5</i>	153
9.3.	Comparaison entre PrOALDoF-Web et les outils de création des cours.....	154
9.3.1.	<i>Les éléments de la comparaison</i>	154
9.3.2.	<i>Les méthodes de création des cours</i>	155
9.3.2.1.	Méthode Manuelle.....	155
9.3.2.2.	Méthode Semi-automatique.....	155
9.3.2.3.	Méthode Automatique.....	156
9.3.3.	<i>Les résultats de la comparaison</i>	156
9.3.3.1.	Méthode de création.....	158
9.3.3.2.	Méthode d'enrichissement.....	158
9.3.3.3.	Facilité d'utilisation.....	159
9.3.3.4.	Recherche du contenu.....	159
9.3.3.5.	Expérience des auteurs.....	159
9.3.3.6.	Qualité du contenu.....	160
9.3.3.7.	L'alignement du but d'apprentissage.....	160
9.3.3.8.	Feedback et adaptation.....	161
9.3.3.9.	Motivation.....	161
9.3.3.10.	Présentation du design.....	162
9.3.3.11.	Ergonomie de l'interaction.....	162
9.3.3.12.	Accessibilité.....	163
9.3.3.13.	Réutilisation.....	163
9.3.3.14.	Conformité des normes standards du compliment.....	163
9.4.	Conclusion.....	164
10.	Conclusion générale.....	166

11. Perspectives	170
PARTIE - III : ANNEXES.....	171
12. L'environnement et les outils utilisés	172
12.1. Introduction.....	172
12.2. L'environnement Java : NetBeans	172
12.2.1. <i>Présentation</i>	172
12.2.2. <i>Historique</i>	172
12.2.3. <i>Installation</i>	173
12.3. Protéger2000.....	173
12.3.1. <i>Présentation</i>	173
12.3.2. <i>Les différentes fonctionnalités de Protégé 2000</i>	175
12.3.3. <i>Installation</i>	175
12.4. Api Google.....	175
12.4.1. <i>Définition</i>	175
12.4.2. <i>Description</i>	176
12.4.3. <i>Les étapes pour utiliser l'API Google</i>	176
12.4.4. <i>Fonctionnement</i>	177
12.4.5. <i>Liste des API Google</i>	177
12.5. Api Excel	178
12.5.1. <i>Description</i>	178
12.5.2. <i>Quelques caractéristiques</i>	178
12.6. Api JENA	179
12.6.1. <i>Description</i>	179
12.6.2. <i>Les composantes de JENA</i>	179
Références	180
Références Personnelles	191
Liste des Abréviations	192

Table des Illustrations

Liste des Figures

Figure 1	Plan du mémoire	5
Figure 2	Principe du Fonctionnement des outils de recherche	18
Figure 3	Les différents modules d'une plate-forme classique de E-Learning	46
Figure 4	Les couches du Web Sémantique.....	52
Figure 5	Le cycle de vie de l'Ontologie	67
Figure 6	L'architecture générale de PrOALDoF-Web.....	100
Figure 7	Description textuelle de l'algorithme PrOALDoF-Web	102
Figure 8	Etapes de la Recherche dans PrOALDoF-Web	108
Figure 9	Exemple de recherche personnalisée (Word) dans Google	112
Figure 10	Exemple de recherche personnalisée (PDF) dans Google	113
Figure 11	Extrait hiérarchique du 'Squelette Humain'	117
Figure 12	Un extrait de l'ontologie 'Squelette Humain' sous Protégé2000.....	118
Figure 13	Un extrait graphique de l'ontologie de domaine créée 'Squelette Humain' .	120
Figure 14	Extrait d'une ontologie sous Protégé2000	122
Figure 15	La transformation de l'ontologie sous le format Excel	123
Figure 16	Extrait d'un document PDF segmenté.....	125
Figure 17	Le principe du processus de fusion de trois ressources Web.....	137
Figure 18	Fusion des sous fragments	138
Figure 19	Extrait graphique de l'ontologie Pédagogique.....	141
Figure 20	Extrait graphique de la hiérarchie de l'ontologie Pédagogique	142
Figure 21	L'approche d'Adaptation basée sur la fusion des sous fragments.....	144
Figure 22	Diagramme des différentes méthodes de création des cours	165
Figure 23	La page principale de NetBeans IDE 6.5.....	173
Figure 24	Projet ouvert sous Protégé v 3.2.1	174

Liste des Tableaux

Tableau 1 Les adresses des sites Web	32
Tableau 2 Comparaison entre une formation traditionnelle et le E-Learning	43
Tableau 3 La Hiérarchie du cours à partir de l'ontologie de domaine	122
Tableau 4 Une partie du modèle Excel pour créer la base 'BNF'	133
Tableau 5 La sauvegarde de la méthode 'DBRS' dans le modèle Excel	133
Tableau 6 Un extrait du modèle Excel d'une portion de la partie Cou et Tronc	135
Tableau 7 Les résultats obtenus en appliquant la méthode 'DBRS'	152
Tableau 8 Eléments de comparaison dans LORI 1.5	154
Tableau 9 Eléments de comparaison proposés dans PrOALDoF-Web	155
Tableau 10 Les niveaux de comparaison	157
Tableau 11 Résultats de la comparaison	157

Liste des Formules

Formule 1 Degré de Pertinence 'DP'	126
Formule 2 Degré de Pertinence Moyen 'DPM'	128
Formule 3 Distance Basée sur les Règles Sémantiques (Moyenne) 'DBRS _M '	129
Formule 4 Distance Basée sur les Règles Sémantiques 'DBRS _D '	130
Formule 5 Précision d'une association 'PA' 'DBRS _{PA} '	131

1. Introduction générale

1.1. Contexte et cadre de la recherche

Les nouvelles Technologies de l'Information et de la Communication 'TIC' améliorent profondément nos façons de nous informer, de communiquer et de nous former. Cette émergence technologique fait apparaître un nouveau mode d'apprentissage connu sous le nom de E-Learning. Celui-ci est basé sur l'accès à des formations en ligne, interactives et parfois personnalisées, diffusées par l'intermédiaire d'un réseau (Internet ou Intranet) ou d'un autre média électronique. Cet accès permet de développer les compétences des apprenants, tout en rendant le processus d'apprentissage indépendant du temps et du lieu.

Il peut être observé que les enseignants sont des acteurs qui ont vu leurs rôles s'enrichir par le développement de compétences visant à tirer profit des avantages des 'TIC'. Nous sommes devant un processus complexe qui requiert de la part des enseignants, une adaptation à des environnements numériques, une forme de travail de plus en plus collaborative et une adoption de méthodes favorisant l'apprentissage actif des apprenants.

Dans ce scénario, les méthodes de création de contenus éducatifs ont toujours posé un accent particulier sur la transformation de matériels didactiques du format analogique en format numérique ; des ressources imprimées en documents 'scannés', transcrits à l'aide d'éditeurs de textes permettant leur distribution sous format PDF ou pages HTML. Ce type d'usage se conjugue le plus souvent avec l'emploi de plateformes à distance ou *Learning Management Systems* (LMS), qui offrent la possibilité d'écrire à l'écran, de saisir du texte et d'importer des médias directement sur un navigateur Web qui s'occupe de leur traitement et de leur mise en œuvre selon les profils des apprenants.

Avec l'avènement de l'Internet, la très grande masse d'informations devenue disponible nécessite aujourd'hui de consacrer une partie considérable de notre temps à l'extraction de l'information pertinente. Dans un espace comme le Web, l'un des enjeux concernant l'éducation réside dans la création des contenus pédagogiques. Les enseignants sont confrontés à la nécessité de produire ou de numériser des matériaux éducatifs non seulement pour les présenter aux apprenants, mais aussi pour les rendre disponibles aux autres enseignants tout en utilisant des stratégies pertinentes pour les découvrir, les partager, les stocker et enfin les utiliser et réutiliser de multiples façons.

L'enseignant ou bien le constructeur du cours équipé seulement d'outils de recherche d'information sur les réseaux tels que WWW, ne peut pas faire face au flux d'information générée. Au lieu de laisser l'enseignant dépenser son temps à chercher l'information dont il a besoin, la tendance actuelle est de concevoir des mécanismes qui permettent de lui faciliter la tâche en lui faisant parvenir l'information qui l'intéresse.

De nombreux documents, et plus généralement de nombreuses ressources, peuvent être utilisées dans le cadre d'une formation en ligne (ou E-Learning). Certaines sont produites 'en interne' par les différents acteurs impliqués dans la formation, d'autres sont disponibles sur le web : cours en ligne, supports de cours, transparents, supports de présentations orales, bibliographies, foires aux questions, notes de lectures, etc.

Et avec l'avènement des ontologies, les enseignants créent des bases de cours (domaine d'apprentissage) en se basant sur plusieurs ontologies, par exemple une ontologie de domaine décrivant les concepts du domaine de la formation ; une autre pédagogique décrivant l'ensemble des notions utiles à une formation particulière ; et nous pouvons intégrer plusieurs ontologies ; tout en assurant l'objectif de chaque ontologie intégrée dans la formation. Le but de cette démarche est de construire le cours dynamiquement via les concepts existants et toutes les ressources disponibles dans la base du système de la formation. Le comportement de l'utilisateur (son état actuel, ses connaissances, son état psychologique) joue un rôle très important dans cette construction dynamique.

Une telle démarche manuelle pour trouver les briques d'informations nécessaires dans la formation est répétitive et coûteuse en temps. Du fait que, les informations pertinentes sont très éparpillées dans le Web, ainsi que les documents sont nombreux et peuvent avoir une grande taille en terme nombre de mots, ce qui demande un effort pour la lecture et la compréhension de ces informations. De plus, les index des pages Web des moteurs de recherche sont en constante mise à jour.

1.2.Objectif et approche

Notre problématique de recherche consiste à améliorer la construction manuelle vers une construction automatique. Donc, dans notre projet, nous avons proposé une approche pour créer un domaine d'apprentissage sans aucune intervention des concepteurs (enseignants).

L'approche commence par : trouver les documents les plus pertinents qui se trouvent sur le Web, les télécharger, les filtrés en se basant sur une ontologie de domaine et sur les règles sémantiques, enfin, extraire les parties les plus pertinentes en appliquant la méthode de fusion proposée afin de créer le domaine d'apprentissage de l'approche proposée. Par la suite, une adaptation basée fusion sera effectuée pour créer les cours dynamiquement à partir du domaine d'apprentissage créé, dont le but, obtenir un cours adaptatif au profil de l'apprenant.

Notre objectif à travers cette approche proposée est de réduire l'immense espace du Web qui contient des milliards de pages Web. Cet espace devient personnalisé avec l'adaptation directe des cours aux profils des apprenants afin de répondre à leurs besoins et afin d'augmenter leurs satisfaction en fournissant une bonne formation extensible à tout changement et mise à jour, mais avec des ressources éducatives fiables.

Dans le cadre de notre projet nous proposons de construire les ressources impliquées dans la formation de façon automatique à partir du Web, donc faciliter la tâche des auteurs (enseignants), tout en leur laissant la possibilité d'analyser les ressources produites pour apporter des modifications (rectifier des erreurs commises, ajouter d'autres informations,...etc.) en intervenant les experts impliqués dans le processus de construction des cours de la formation.

Une formation est constituée d'acteurs (apprenants, enseignants, etc.), de différentes ressources (définitions, exercices, ...etc.) rédigées sous différentes formes (livres, sites web, ...etc.) et sur différents supports (papier, vidéo, audio, etc.), notre projet vise à aider les enseignants pour construire les différentes ressources textuelles d'une façon automatique, afin de réduire le nombre d'enseignants impliqués dans la formation et de trouver les bonnes parties, c'est-à-dire, les plus pertinentes pour chaque ressource. Donc, les cours dans la phase finale seront construits via plusieurs fragments fusionnés. Ces derniers peuvent être adaptés directement aux profils des apprenants en se basant sur leurs niveaux.

PrOALDoF-Web est un projet de recherche qui consiste à réaliser un système d'enseignement en ligne dont le contenu pédagogique est adaptatif au profil de l'apprenant et dont la source d'information disponible sur le Web est bien entendu extraite et réorganisée selon un canevas prédéfini du domaine d'enseignement afin d'enrichir convenablement les connaissances de l'apprenant selon ses aptitudes et selon son niveau.

PrOALDoF-Web est une abréviation en anglais “**Proposition and Organization of an Adaptive Learning Domain based on Fusion from the Web**”. Donc, PrOALDoF-Web propose une nouvelle approche et organise un domaine d’apprentissage adaptatif au profil de l’apprenant en se basant sur la méthode de fusion des ressources disponibles dans le Web.

Pour répondre aux objectifs de la thèse, nous donnons notre approche :

- La formulation d’un corpus de documents qui joue le rôle d’un support physique de test (comme un moteur de recherche), vu la limite des requêtes et le nombre de résultats par jour des moteurs de recherche qui proposent des API gratuites ;
- La proposition d’une méthode de recherche des documents les plus pertinents (plus précisément, les morceaux documentaires les plus pertinents) en se basant sur une ontologie de domaine ;
- La proposition d’une méthode de fusion de morceaux de texte extraits à partir du corpus ;
- L’adaptation du contenu pédagogique en utilisant la méthode de fusion et une ontologie pédagogique.

PrOALDoF-Web est composé de trois parties : 1) Partie Recherche, 2) Partie Fusion, 3) Partie Adaptation. Les trois parties peuvent travailler simultanément ou individuellement, mais à la fin du processus de traitement les trois parties seront complémentaires.

A travers PrOALDoF-Web nous développons une nouvelle approche dans les systèmes éducatifs en ligne adaptatifs, qui peut être intégrée à n'importe quel LMS.

Le contenu d'apprentissage peut être réutilisé et partagé entre les différentes plates-formes. La source d'information de contenu est le Web. La méthode de recherche et de filtrage est basée sur une ontologie de domaine et sur les règles sémantiques. L'approche d'adaptation est basée sur la méthode de fusion via l'ontologie de domaine décrivant les concepts du cours et l'ontologie pédagogique décrivant la technique d'adaptation.

Ce qui est nouveau dans notre travail est la fusion de plusieurs sous fragments des ressources Web jugées pertinentes. En conséquence, l’amélioration de la qualité du contenu de la formation et l’obtention d’un système éducatif adaptatif, fiable, riche et dynamique; dynamique est dans le sens d'enrichissement et de mise à jour.

1.3. Plan du mémoire

Le mémoire de la thèse est composé de deux grandes parties : Partie I (Etat de l’art) et Partie II (Notre projet PrOALDoF-Web) (Voir Figure 1 ci-dessous).

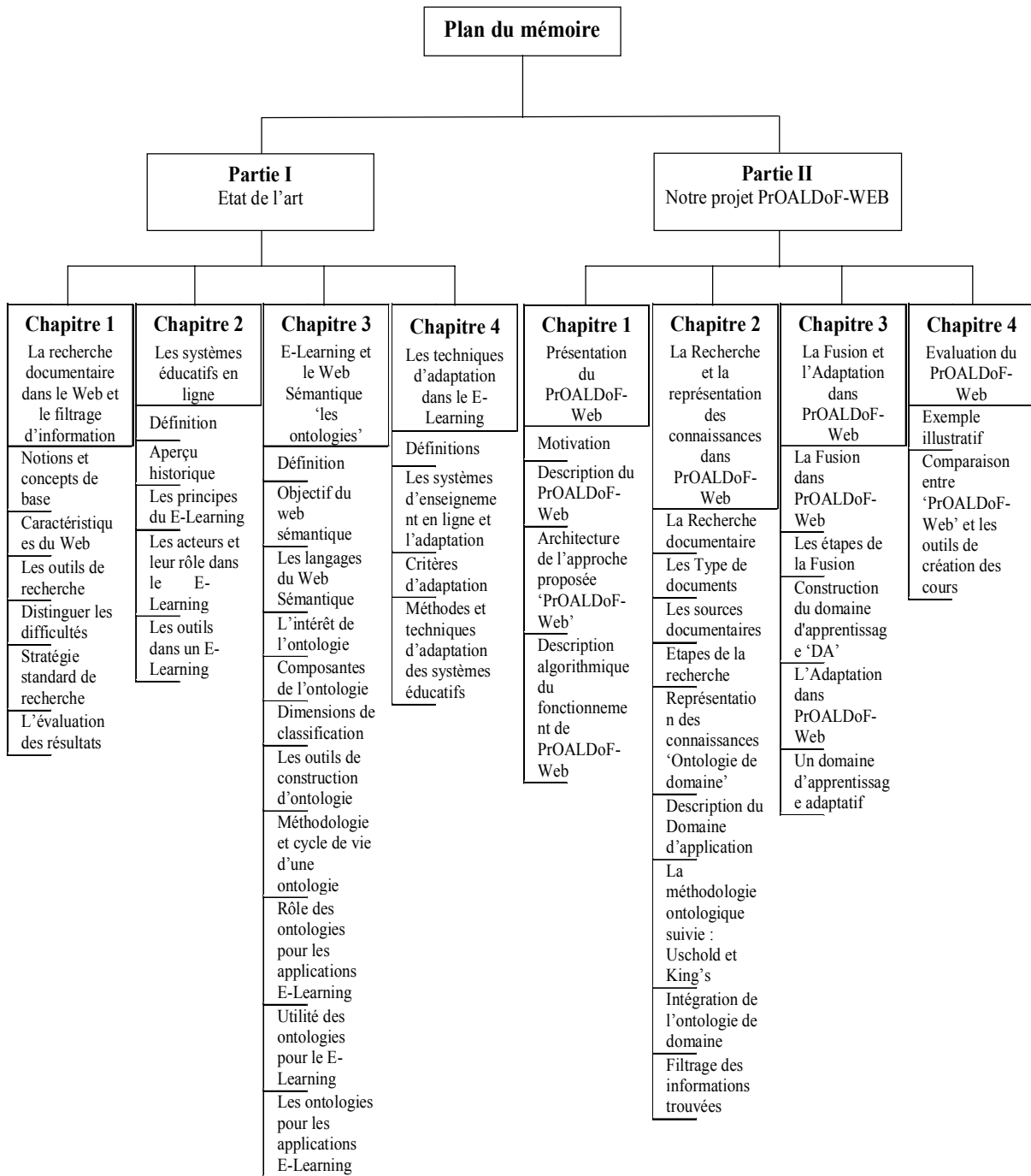


Figure 1 Plan du mémoire

La première partie ‘I’ contient quatre chapitres comme suit :

Le premier chapitre se focalise sur la recherche documentaire dans le Web et le filtrage d’information. Nous présentons dans ce chapitre les notions et les concepts de base de la recherche d’informations dans le Web et le filtrage des résultats trouvés. Nous donnons dans un premier lieu les caractéristiques du Web, puis nous présentons quelques difficultés rencontrées et nous proposons des solutions pour éviter la perte du temps et d’effort fournissent par les internautes en général et par les apprenants et les enseignants en spécial. Par la suite, nous détaillons une méthodologie de recherche à suivre afin d’obtenir les meilleurs résultats.

Le deuxième chapitre est sur les systèmes éducatifs en ligne. Dans cette partie, nous discutons les caractéristiques générales d’un E-Learning en détaillant les principes de base de ce concept. Nous plongeons ensuite au cœur du E-Learning en décrivant les différents éléments qui le composent, les acteurs qui l’animent. Enfin, les outils qui permettent de gérer ces formations.

Le troisième chapitre étudie le E-Learning et le Web sémantique ‘les ontologies’. Le chapitre définit le web sémantique et les ontologies, puis il décrit les objectifs de la nouvelle génération du Web, son architecture globale, ses langages candidats, à savoir : XML, RDF, RDFS, OWL, Topic Maps. Afin de bien argumenter l’utilité des ontologies, nous décrivons les composantes d’une ontologie. Puis, nous discutons l’apport des ontologies avec la représentation des connaissances. Après, nous présentons les différentes typologies, les outils de construction, les méthodologies et le cycle de vie d’une ontologie. Enfin, nous discutons le rôle joué par les ontologies dans les applications E-Learning.

Le dernier chapitre dans la première partie est sur les techniques d’adaptation dans le E-Learning. Le chapitre décrit les notions de base de l’adaptation. Ensuite, il présente l’apport dans les systèmes éducatifs en ligne intégrant des dispositifs d’adaptation. Ces systèmes sont capables d’adapter selon les critères suivants: Les données de l’utilisateur, les données d’usage, et l’environnement. Puis, il donne une vision globale sur les paramètres d’adaptation qui sont relatifs à l’utilisateur et ceux qui sont indépendants, les différentes méthodes et techniques d’adaptation et ses éléments.

La deuxième partie ‘II’ est composée de quatre chapitres comme suit qui contiennent notre contribution de recherche et les différents concepts liés à cette contribution :

Le premier chapitre nommé : La présentation du projet PrOALDoF-Web. Nous commençons par une motivation, puis une description du projet et ses objectifs, par la suite nous détaillons l'architecture générale du projet avec ses différentes parties. Enfin, nous donnons l'algorithme de l'approche proposée avec le principe du fonctionnement de tout le projet.

Le deuxième chapitre est sur la recherche et la représentation des connaissances dans PrOALDoF-Web. Ce chapitre présente les différentes étapes de la recherche dans le projet PrOALDoF-Web en commençant par une définition du terme recherche documentaire. Ensuite, nous citons les types des documents. Puis, nous parlons sur les sources documentaires. Les étapes de notre recherche sont détaillées dans le point suivant en commençons par la formulation automatique de la requête de recherche, trouver les informations pertinentes et enfin analyser le contenu des informations trouvées. Puis, une représentation des connaissances 'Ontologie de domaine' est définie. L'élément suivant, est bien entendu le domaine d'application qui est dans notre cas le squelette humain dans lequel nous avons créé notre ontologie de domaine et notre ontologie pédagogique. La méthodologie ontologique Uschold et King's est celle utilisée pour créer nos ontologies qui sont intégrées via l'API JENA. Puis, nous détaillons notre méthode de filtrage pour extraire les informations les plus pertinentes. Enfin, nous avons discuté les éléments de l'évaluation des résultats trouvés par la recherche Web afin de juger leurs pertinences et leurs fiabilités.

Le troisième chapitre parle de la fusion et l'adaptation dans PrOALDoF-Web. Nous voulons par ce chapitre, la présentation de nos deux approches essentielles dans notre projet PrOALDoF-Web. La première approche est la fusion des ressources Web jugées pertinentes. L'approche utilise un modèle proposé et des méthodes statistiques et sémantiques pour sauvegarder tous les résultats trouvés. Nous détaillons le principe de la fusion avec ses différentes étapes afin d'obtenir le but fixé. La deuxième approche est l'adaptation des ressources pertinentes aux profils des apprenants en se basant sur une ontologie pédagogique dont le but est de construire un domaine d'apprentissage adaptatif.

Le dernier chapitre dans la deuxième partie est sur l'évaluation du PrOALDoF-Web. A travers ce chapitre, nous présentons dans un premier lieu un exemple illustrant tout le processus essentiel dans notre projet qui est la fusion et l'adaptation, car la recherche est un processus à l'intérieur du système que nous ne pouvons pas le visualisé, mais, nous avons le simulé via la recherche manuelle de Google. Par la suite, nous comparons 'PrOALDoF-Web' avec deux types d'outils de création des cours existants : commercial ou éducatif.

PARTIE – I : ETAT DE L'ART

Contenu

Chapitre I : *La recherche documentaire dans le Web et le filtrage d'information*

Chapitre II : *Les systèmes éducatifs en ligne*

Chapitre III : *E-Learning et le Web Sémantique 'les ontologies'*

Chapitre VI : *Les techniques d'adaptation dans le E-Learning*

2. La recherche documentaire dans le Web et le filtrage d'information

'Trouver l'information est un art, pas une science'

Jean-Pierre Lardy

Nous présentons dans ce chapitre les notions et les concepts de base de la recherche d'informations dans le Web et le filtrage des résultats trouvés. Nous donnons dans un premier lieu les caractéristiques du Web, puis nous présentons quelques difficultés rencontrées et nous proposons des solutions pour éviter la perte du temps et d'effort fournissent par les internautes en général et par les apprenants et les enseignants en spécial. Par la suite, nous détaillons une méthodologie de recherche à suivre afin d'obtenir les meilleurs résultats.

2.1.Introduction

Le Web n'est qu'une source d'information parmi beaucoup d'autres. Il peut se révéler très utile dans certains cas, et presque complètement inutile dans d'autres, malgré qu'il présente la plus grande bibliothèque du monde avec des milliards de pages sur le réseau internet (entreprises, gouvernements, associations, universités, pages personnelles...). Mais, le problème majeur devant les utilisateurs c'est comment retrouver les informations dont ils ont besoin dans cette immensité. Nous constatons que les utilisateurs passent en moyenne 80 % de leur temps à rechercher de l'information et 20 % à la traiter.

Il y a des outils qui vous permettent de consulter et de rechercher de l'information. Ces outils examinent les nombreuses pages présentes sur le Web et mettent à votre disposition d'absolus indexes. Donc, retrouver l'information dont vous avez besoin peut s'avérer une action difficile malgré que ces indexes ne couvrent qu'une minime partie du web. La difficulté est dans le choix d'un outil qui se fait dans la plus part du temps un peu au hasard (un dont nous avons entendu parler, un dont le nom est facile à mémoriser) et rarement suite à une étude comparative des qualités inhérentes de chaque outil. Une autre difficulté dans la tâche de filtrage du résultat obtenu (un immense nombre de pages trouvées).

Il y'a des travaux et des sites qui offrent la possibilité de création des sites personnalisés (type du site, thèmes, couleurs, ...etc.), mais, dans notre travail nous intéressons à la personnalisation du contenu des sites en facilitant la tâche de la recherche surtout dans la formulation de la requête, puis le filtrage du résultat.

2.2. Notions et concepts de base

2.2.1. Recherche d'informations vs Recherche documentaire

Selon les définitions de l'AFNOR et (Dinet, Rouet et Passerault, 1998), il est nécessaire de distinguer recherche d'informations et recherche documentaire. La recherche d'informations correspond à un ensemble d'actions, méthodes et procédures ayant pour objet d'extraire d'un ensemble de documents les informations voulues (AFNOR, 1993, p. 99), alors que la recherche documentaire correspond à l'ensemble d'actions, méthodes et procédures ayant pour objet de retrouver dans des fonds documentaires les références des documents pertinents (AFNOR, 1993, p. 99). En d'autres termes, effectuer une recherche documentaire équivaut à identifier et à accéder à des ressources informationnelles qui ont déjà été traitées et éditées (Bibeau, 1998).

Une autre définition de la recherche documentaire est qu'elle est une démarche systématique, qui consiste à identifier, récupérer et traiter des éléments divers (chiffres, bibliographie, textes...) sur un sujet donné. Cette identification des informations est une étape indispensable à toute synthèse des connaissances et revue de la littérature dans le domaine médical, comme dans d'autres domaines. Cette démarche doit être la plus pertinente possible et tendre vers l'exhaustivité [(IRDES, 2012) et (UVMaF, 2011)]. Elle nécessite donc :

- une parfaite connaissance des multiples sources d'information ;
- la maîtrise des outils et des stratégies de recherche.

En conséquence, nous pouvons dire que la recherche documentaire est une activité compliquée entre la résolution de problème et la prise de décision.

2.2.2. Filtrage des informations dans le Web

Est un ensemble de techniques et de méthodes mises en place dans le but de limiter l'accès à certains sites normalement accessibles sur le réseau Internet. Ces sites sont nommés indésirables. De plus il s'agit d'un processus consistant à extraire des informations dites "pertinentes" et de qualité à partir d'une grande masse d'informations. L'objectif est de repérer et d'extraire des données susceptibles de correspondre aux besoins et intérêts indiqués par l'utilisateur en formulant leurs requêtes, les données peuvent être extraites de sources différentes évoluant dans le temps [(WIKIVERSITY, 2013), (Dessonnaz, 2013), (WIKIPEDIA, 2013) et (Allan et al, 2002)].

Nous pouvons citer quelques manières de cette technique :

- L'accès restreint : l'URL du site fait partie d'une liste noire ;
- Filtrage du contenu : analyse basée mots clés ;
- Type du document : PDF, DOC, XLS ...etc. ;
- Nature du contenu : Jeux, Ouvrages, Téléchargements ...etc. ;
- Type de la ressource : Texte, Image, Vidéo ...etc.

2.2.3. Annuaire

Un annuaire est un répertoire de sites, organisé selon un classement thématique ou géographique fait par un éditorialiste. Chaque site est commenté par l'éditorialiste et classé par catégorie. L'annuaire est surtout utile pour explorer un sujet général, trouver des sites ressources dans un domaine ou encore trouver des sites similaires sur un même thème. (IAAT, 2005).

Les annuaires recensent selon une arborescence souvent complexe de catégories et de sous-catégories, les sites Internet qui leur sont soumis (par leurs auteurs ou leurs propriétaires, le plus souvent) ; leur fonctionnement est proche de celui des Pages Jaunes : ils reposent sur une déclaration préalable, associée le plus souvent à une validation qui peut faire intervenir de façon aléatoire ou systématique un agent humain. Nous pouvons citer comme exemple d'annuaire: **Yahoo!**, **Voila**, **Nomade** ou encore **Lycos**. Le nombre de sites référencés est inférieur à ce que proposent en général les moteurs de recherche, et la pertinence réelle des sites référencés est difficile à évaluer. Ces annuaires, d'abord portés par leur inscription dans un portail offrant d'autres services (voir Yahoo!), sont de moins en moins utilisés – l'inscription payante, permettant un référencement plus rapide et plus complet, disparaît de la plupart d'entre eux ; sur certains, c'est la page d'inscription elle-même qui est devenue introuvable [(Grari, 2007), (AEC, 2005), (IAAT, 2005) et (Pouilly, 2007)].

2.2.4. Moteur de recherche

Un moteur de recherche est un outil de recherche de pages web. A la différence de l'annuaire, c'est un répertoire de pages et non de sites. Le moteur de recherche balaye le réseau internet sur un mot ou une expression et affiche les pages web, les fichiers (Excel, Word, pdf, PowerPoint...etc.) traitant le sujet demandé. Il peut également répertorier des groupes de discussion, des images, des actualités... etc. (IAAT, 2005).

Les moteurs de recherche fonctionnent différemment : ils commencent par constituer un index de pages Internet (et non de sites) ; les auteurs peuvent soumettre leurs pages pour inscription dans cet index, mais la grande masse en est récupérée de façon automatique, à l'aide d'un robot (un logiciel) qui suit et analyse les liens rencontrés dans les pages et navigue ainsi de l'une à l'autre.

Une copie de ces pages est stockée sur des serveurs dédiés (le chiffre exact est tenu secret, mais nous estimons dans le cas de Google à plus de 15 000 le nombre de PC permettant de faire tourner ce 'cache') [(AEC, 2005) et (Grari, 2007)].

Lorsque l'internaute tape ensuite un mot dans la zone de saisie, un programme l'analyse sommairement (reconnaissance de la langue d'origine, filtre orthographique), puis parcourt les pages ainsi stockées à sa recherche. Ce sont les résultats de cette recherche, classés selon une méthode que nous discuterons plus loin, qui sont présentés à l'internaute, avec l'adresse de la page correspondante ainsi qu'un extrait du contexte dans lequel le mot apparaît au sein de la page.

La quantité de pages examinées est beaucoup plus importante que dans le cas des annuaires (Google, jamais en peine de surenchère en la matière, avec un chiffre de plus de 10 milliards de pages, Microsoft et Yahoo! sont dans les mêmes ordres de grandeur) ; la recherche se fonde pour l'essentiel sur le contenu des pages, et non sur le résumé d'un site fourni par son auteur [(AEC, 2005) et (Grari, 2007)].

L'écrasante suprématie de Google et de quelques grands acteurs émergents comme Yahoo! (qui a renoncé à utiliser la technologie Google pour tenter d'imposer la sienne propre) ou comme Microsoft avec MSN search ne doit pas conduire à négliger d'autres moteurs généralistes comme le français Exalead, Altavista (en nette perte de vitesse alors qu'il dominait outrageusement le secteur aux temps héroïques du début du Web grand public), Hotbot ou encore Reacteur – ni des moteurs innovants comme Blinkx ou spécialisés (par zone géographique, par thème, par type de sources indexées, etc.) (AEC, 2005).

2.2.5. Méta-moteurs

Les méta-moteurs, quant à eux, sont des outils logiciels qui interrogent en simultanément plusieurs moteurs de recherche ; principal défaut de la plupart d'entre eux, Google, acteur technologique incontournable, interdit l'accès à son moteur à ce type d'outils.

Quant aux agents ' intelligents ', il s'agit d'outils paramétrables destinés à effectuer de façon semi-autonome des missions de recherche d'information et surtout de veille (observation des changements) sur différentes sources choisies [(AEC, 2005) et (Grari, 2007)].

2.2.6. Moteur vs annuaire

Il y a quelques années, certains sites étaient spécialisés en tant que moteur de recherche Web et d'autres en tant qu'annuaire ou répertoire (*directory*). Maintenant, les sites de moteurs de recherche les plus connus offrent les deux types de fonctionnement (et même davantage). Par exemple Google était un moteur de recherche et Yahoo était un annuaire : l'un et l'autre site offrent maintenant les deux modes de fonctionnement (Rézeau, 2006).

2.2.7. L'exploration ou Crawl

Le web est systématiquement exploré par un robot (crawler, spider) suivant tous les hyperliens et récupérant les ressources jugées intéressantes (MIC, 2007).

2.2.8. Action de recherche

Correspond à la partie requêtes du moteur, qui restitue les résultats. Un algorithme est appliqué afin de pouvoir présenter les résultats des recherches selon un ordre de pertinence. L'algorithme tient généralement compte du contexte du mot clé (titre, paragraphe, hyperlien...) et de la ressource (ressources liées, popularité du site...) (MIC, 2007).

2.2.9. L'indexation

Consiste à extraire et à enregistrer dans un gigantesque index les mots considérés comme significatifs de chaque ressource récupérée. Le moteur retrouve ainsi plus rapidement où se situe un terme significatif donné. Les termes non significatifs s'appellent des mots vides (' de ', ' le ', ' un '...) (MIC, 2007).

2.2.10. Banque des données

La banque de données est :

- produite par une institution, une société commerciale ou un service de documentation ;
- stockée sur un serveur ;
- et affichée sur le micro-ordinateur de l'utilisateur en réponse à une requête.

Différentes natures de l'information sont prises en compte dans les banques de données : références bibliographiques, texte intégral, bases de connaissances (données factuelles ou numériques), images...etc. [(UVMaF, 2011) et (IRDES, 2008)].

2.2.10.1. Les bases bibliographiques

Décrivent, sous forme de références bibliographiques (ou notices), divers documents (articles, ouvrages, rapports, congrès...), avec un apport d'indexation (descripteurs) et éventuellement d'analyse (résumé). Une notice bibliographique constitue la carte d'identité du document : auteur(s), titre, éditeur, date de publication, mots-clés, résumé... etc.

- Un thesaurus est un vocabulaire structuré de mots où chaque terme (descripteurs ou mots-clefs ont entre eux des relations sémantiques (hiérarchie, synonymie, association)).

2.2.10.2. Les bases textuelles

Le texte intégral est disponible en banques de données dans de nombreux domaines : juridique, scientifique et technique, et pour la presse. Le texte peut être codé, les termes sont alors interrogeables, ce qui est le cas dans l'accès en ligne, ou bien le texte est numérisé, c'est-à-dire visible en tant qu'image mais non interrogeable directement.

2.2.10.3. Les bases de connaissance (factuelles ou numériques)

Fournissent des renseignements directement et immédiatement exploitables par les utilisateurs (par exemple : les signes d'une maladie, les propriétés d'un médicament, des guidelines ou référentiels de pratique clinique).

Les bases informatiques sont devenues le principal outil de la recherche bibliographique. Seul le traitement informatique permet la gestion du volume toujours croissant d'informations : rapidité, exhaustivité, pertinence.

Elles proposent à l'utilisateur des notices bibliographiques ;

- Titre, auteurs, revue, année, volume...
- Résumé (abstract)
- Mots-clés (ou descripteurs) extraits d'un thesaurus.

2.3. Caractéristiques du Web

2.3.1. Le Web est énorme

La difficulté provient en partie de la taille du Web et du nombre considérable de documents qui s'y trouvent hébergés, répartis sur de nombreux sites. Il est impossible de connaître le nombre exact de ces sites et de ces documents, qui augmentent de façon exponentielle [(Rézeau, 2006) et (BULCO, 2012)].

Voici quelques chiffres valables en jan. 2013 (Bonnet, 2013) :

- 2,4 milliards : le nombre d'utilisateurs d'internet dans le monde (dont 1,1 milliard en Asie et 565 millions en Chine)
- 634 millions : le nombre de sites internet en activité, soit 51 millions de plus par rapport à 2011
- 144 milliards : le nombre d'e-mails envoyés chaque jour en moyenne
- 68,8% : la part des spams dans ces e-mails (50,76% pour des articles de 'pharmacie')
- Google, roi des moteurs de recherche avec 1 200 milliards recherche en 2012 (part de marché de 67%)
- YouTube affole les compteurs avec 4 milliards heures de vidéos vues chaque mois en moyenne sur et un 1 milliard vues enregistré par Gangnam Style du coréen Psy, première vidéo à accomplir cette performance
- Facebook, premier sur les réseaux sociaux avec 1 milliard utilisateurs actifs de Facebook (chiffre atteint en octobre 2012) et 300 millions photos ajoutés chaque jour en moyenne.
- Le mobile gagne du terrain avec 13% comme part de l'internet mobile dans le trafic mondial (4% en 2010).

2.3.2. Le Web vs Bibliothèque

- On entend parfois dire que le Web est *une immense encyclopédie* ou *une immense bibliothèque*. Cette image est **fausse**, tout simplement parce que les documents présents sur le Web ne sont pas indexés ni classés. Une image plus exacte du Web serait celle d'une immense bibliothèque dont tous les ouvrages auraient été éparpillés au hasard en un tas informe (et pour un bon nombre de ces ouvrages les pages elles-mêmes auraient été arrachées et dispersées !). [(Rézeau, 2006) et (BULCO, 2012)].

- Lorsque vous dites que vous “faites une recherche sur le Web”, en réalité vous ne faites pas cette recherche sur le Web lui-même. En effet, il n'est pas possible de faire une recherche sur ou dans le Web directement. Le Web est la totalité des innombrables pages Web qui sont hébergées sur des ordinateurs (ou serveurs) localisés dans le monde entier. Votre navigateur ne peut pas aller fouiller tous ces ordinateurs ou serveurs un par un de par le monde [(Rézeau, 2006) et (BULCO, 2012)].
- Ce que vous pouvez faire avec votre navigateur, en revanche, c'est accéder aux outils de recherche intermédiaires qui sont disponibles. Vous faites votre recherche dans la base de données ou la collection de sites Web indexés par un moteur de recherche. Cette base n'est qu'une petite partie de la totalité des documents Web existants. Le moteur de recherche répond à votre requête en affichant une liste de liens vers des adresses de pages Web qui semblent y répondre. En cliquant sur les liens vous pouvez accéder aux documents hébergés sur des ordinateurs localisés n'importe où dans le monde [(Rézeau, 2006), (BULCO, 2012) et (Allan et al, 2002)].

2.3.3. Le Web invisible

- Notez par ailleurs qu'une grande partie des documents Web existants ne sont pas accessibles par l'intermédiaire des moteurs de recherche. Tous ces documents font partie de ce que nous appelons le Web invisible [(Rézeau, 2006) et (BULCO, 2011)].
- Les ressources vers lesquels les hyperliens sont créés dynamiquement pour répondre à une interrogation échappent aux robots d'indexation. En effet, un robot n'est pas capable d'émettre des interrogations pertinentes, aucun hyperlien n'est donc créé lors de sa visite.
- Ces ressources constituent de ce qui est parfois appelé le *Web invisible* ou *Web profond*. Il s'agit surtout de ressources provenant de bases de données, accessibles à travers une passerelle. Une étude de la firme Bright Planet a évalué que le Web profond pouvait contenir 500 fois plus de ressources que le Web indexé par les moteurs de recherche [(Rézeau, 2006) et (IRDES, 2012)].

- Le contenu de cette importante source d'informations qu'est le Web est donc d'un volume considérable, mais mal organisé et difficile d'accès. Heureusement nous disposons d'outils spécialisés pour y effectuer nos recherches [(Rézeau, 2006) et (IRDES, 2012)].

2.4. Les outils de recherche

Un outil de recherche est un terme générique pour tout service de recherche d'information sur le Web, combinant bien souvent désormais les procédés d'un répertoire et ceux d'un moteur de recherche, plus parfois de moteurs spécialisés (AERIS, 2006).

Pour trouver un outil précis, avoir une vue d'ensemble des moteurs, des annuaires, chercher un outil spécialisé, ...etc. (Serres, 2011).

Si le monde de l'internet est très mouvant, celui de la recherche de l'information l'est encore plus. La recherche d'information est un domaine en constante évolution et ce à deux niveaux (Lérignier et Marcheix, 2013) :

- Le contenu même du réseau internet change en permanence. Ainsi il n'existe aucun outil de recherche pouvant prétendre indexer l'ensemble des informations disponibles sur internet.
- Le fonctionnement des outils de recherche évolue très régulièrement. Il est donc indispensable de se tenir au courant de ces évolutions.

2.4.1. Fonctionnement des outils de recherche

Le fonctionnement de ces différents outils de recherche reste similaire si l'on considère uniquement les échanges réalisés sur le réseau, permettant d'aboutir au résultat de la recherche (Lérignier et Marcheix, 2013).

Considérons la Figure 2 en page 18. Le principe de fonctionnement de ces échanges est le principe Client Serveur, qui est courant sur les réseaux. Un client se connecte à un serveur, lequel retourne l'information demandée au client.

Dans le cas d'une recherche effectuée sur un outil de recherche et de la consultation des pages de résultat, plusieurs serveurs entrent en jeu.

- D'un côté, le serveur de l'outil de recherche qui stocke le site web permettant d'effectuer la recherche.
- D'un autre côté les différents serveurs qui hébergent les sites web contenant l'information recherchée.

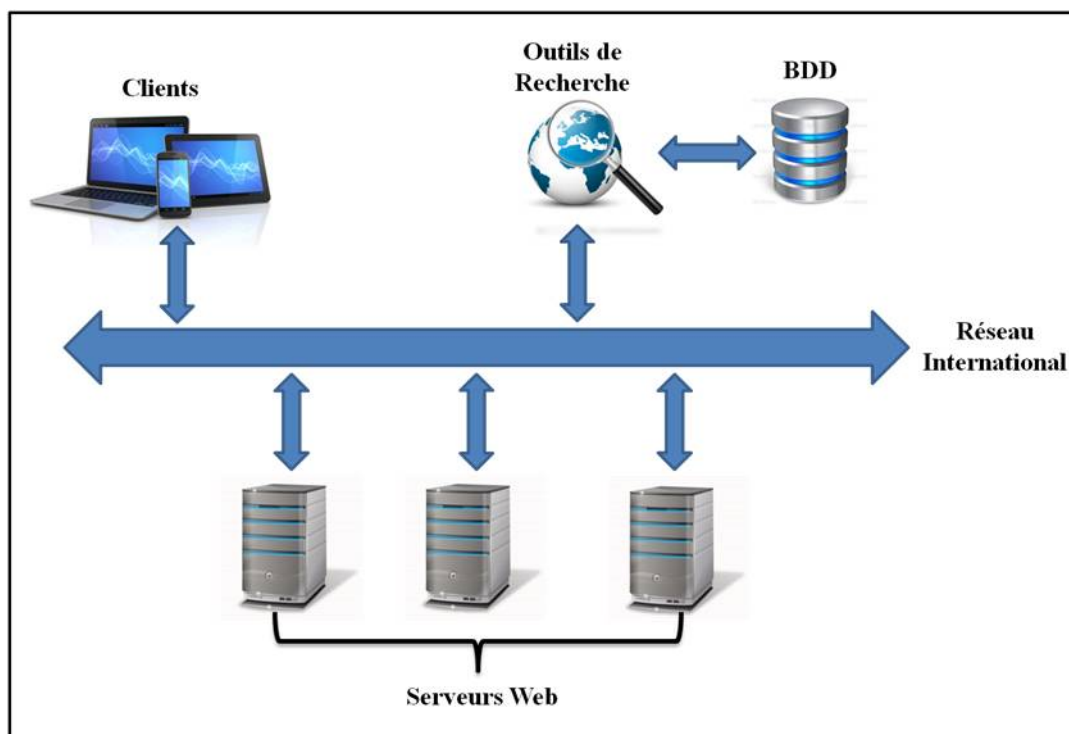


Figure 2 Principe du Fonctionnement des outils de recherche

Le client est la personne qui effectue la recherche. Lors d'une recherche, nous distinguons les étapes successives suivantes :

- Il commence par aller sur le site web de l'outil de recherche. Une requête part donc en direction de ce serveur demandant la page d'accueil de cet outil.
- Le serveur retourne cette page d'accueil au client et le navigateur du client affiche cette page à l'écran.
- Le client tape plusieurs mots clés dans la zone de recherche (par exemple : droit patrimonial) et lance la recherche.
- Une requête part de nouveau vers le serveur de l'outil de recherche.
- La recherche de ces deux mots clés est alors réalisée dans la base de données de l'outil de recherche et non sur les serveurs web qui hébergent effectivement les sites web parlant de droit patrimonial.
- Après interrogation de sa base de données, l'outil de recherche construit une page de résultats contenant des liens vers ces sites web, qu'il retourne au client.

Le client peut alors cliquer sur ces liens hypertexte pour aller consulter sur les serveurs web les sites traitant effectivement de droit patrimonial.

Si le fonctionnement des différents outils de recherche reste similaire au niveau de ces 'échanges réseaux', ce qui les distingue de manière fondamentale c'est la gestion de leur base de données. (Lérignier et Marcheix, 2013).

Afin de pouvoir aboutir à des réponses pertinentes lors d'une recherche, nous trouvons trois types d'outils de recherche sur le Web [(UNJF/C2i, 2010) et (Lérignier et Marcheix, 2013)] : Les outils humains (annuaires, répertoires, guides, listes thématiques), Les outils automatiques que nous les qualifions moteurs de recherche et Les agents intelligents qui permettent d'automatiser la recherche et le traitement des résultats de cette recherche.

2.4.2. Types d'outils de recherche

2.4.2.1. Les outils humains

Ces outils cherchent plus à être qualitatifs qu'exhaustifs dans la gestion de leur base de données (UNJF/C2i, 2010). La caractéristique principale de ces outils est que les sites y sont classés dans une structure arborescente thématique (catégories, sous-catégories, ...etc.) au travers de fiches descriptives créées par des netsurfeurs (personnes, souvent professionnels, en charge de l'évaluation et du référencement de sites dans l'annuaire).

Une fiche descriptive créée et remplie par le netsurfeur comprend en règle générale :

- le titre du site,
- son URL,
- un bref commentaire,
- une liste de mots clés,
- une ou plusieurs notes de qualités permettant d'évaluer le fond, la forme, l'interactivité du site, la facilité de navigation, etc. A ce sujet, la plupart des annuaires récompensent certains sites qu'ils apprécient en leur allouant un signe distinctif. Bien que cela soit subjectif, vous pouvez considérer, lorsque vous voyez apparaître ce signe distinctif sur une page de résultat, qu'il s'agit d'une source d'info intéressante : Yahoo France (paire de lunettes), Nomade.fr (palme), Lycos (chien), Open Directory (étoile).

Contrairement à ce que nous le trouvons dans les moteurs de recherche, dans la base de données de ces outils c'est l'ensemble d'un site web qui est référencé de manière globale et non chaque page web de ce site individuellement (une fiche descriptive est créée pour l'ensemble du site).

Un même site peut être classé dans différentes catégories de l'arborescence. Il faut bien comprendre que les recherches sont ici effectuées sur les fiches descriptives et non dans le contenu même des pages web.

Même si cela paraît peu probable, cela implique qu'un mot recherché peut être trouvé dans une fiche descriptive et ne pas apparaître sur le site lui-même. A l'inverse, la fiche descriptive ne contenant que quelques mots, et même si à priori il s'agit des plus pertinents, nous comprenons aisément que la recherche s'effectue sur un ensemble de mots très réducteur par rapport à ce que contient réellement le site web. Mais c'est là tout l'intérêt de ces outils qui utilisent l'intelligence et l'analyse d'être humains (les netsurfeurs) pour synthétiser et extraire l'essentiel du contenu d'un site web. Le bruit (pages web qui n'ont rien à voir avec l'objet de la recherche) est ainsi bien plus limité.

Trois manières de chercher sur ces outils :

- *En descendant les catégories du plus général vers le plus précis.*

Cela offre 2 avantages. Tous les sites de la rubrique répondent logiquement à votre recherche (limite le bruit). Choix plus exhaustif de sites répondants à votre problématique (ex : si vous tapez 'court métrage' vous ne trouverez pas les fiches contenant l'expression 'film court').

- *En tapant des mots clés.*

En règle générale, lorsque des mots clés sont saisis, l'outil propose d'abord les rubriques dont l'énoncée contient le mot demandé, puis les fiches descriptives des sites qui répondent à la demande (c.à.d. dont le titre, le résumé ou l'adresse contient le mot demandé). Attention : sur ce type d'outil l'erreur à ne pas commettre consiste à utiliser des mots clés trop précis (la recherche s'effectue sur les fiches descriptives et il n'y aurait pas de résultats). A chaque fois que nous tapons un mot clés sur ce type d'outil il faut préalablement se poser la question suivante : est-il probable que le netsurfeur ait utilisé ce mot pour synthétiser tout un site dans une fiche descriptive ?

- *En combinant les deux techniques précédentes.*

Nous effectuons initialement une recherche sur quelques mots clé dans le but de localiser la catégorie pertinente. Nous descendons ensuite dans cette catégorie et nous effectuons une recherche par mots clés limitée à cette catégorie et ses sous-catégories. Cette stratégie s'avère la plupart du temps bien plus efficace que les deux précédentes.

Les bases de données de ces outils offrent plusieurs types de couvertures :

- Généraliste (ils référencent tous les sites sans distinction ou dans une langue).
- Thématique (ils référencent tous les sites traitant d'un domaine précis).
- Spécialisée (ils référencent tous les sites contenant un type particulier de données : images, vidéos, news, ftp, listes de diffusion, etc.).

Nous pouvons donner deux exemples des outils humains : Lycos, Nomade, Guide de Voila et AOL (français) et DMOZ-Open Directory, Yahoo et Looksmart (anglais).

2.4.2.2. Les outils automatiques (moteurs de recherche)

Les moteurs de recherche cherchent à être exhaustifs plus que qualitatifs (UNJF/C2i, 2010). Leurs bases de données contiennent des liens vers des milliards de pages Web (environ 8 milliards de pages pour les plus gros - dont 50% en anglais et 5% en français - et les plus gros outils de recherche ne recensent que 40% des documents disponibles en ligne). Vu la taille des bases de données, pour trouver l'information avec ces outils, la recherche doit être très précise et nécessite impérativement l'utilisation d'opérateurs de requête sous peine de recevoir une réponse très volumineuse.

Il n'y a plus d'intervention humaine dans le processus de référencement, ce qui permet d'atteindre de telles quantités. Contrairement aux outils manuels, ils référencent directement le contenu même des pages web et non de manière globale une fiche descriptive contenant une description globale du site.

A la base, dans un moteur de recherche, l'information n'y est pas non plus classée par catégories thématiques comme dans le cas des annuaires. Des évolutions récentes sont toutefois proposées par certains moteurs de recherche (Exalead, etc.) afin de générer une classification automatique des documents à posteriori (cette technique est appelée 'Clustering').

Nous employons le terme de moteur de recherche pour désigner un site web permettant d'effectuer une recherche. Or ceci est un abus de langage, puisqu'au sens véritable du terme il s'agit du programme (appelé aussi robot ou spider) qui remplace l'intervention des navigateurs en scrutant en permanence le Web pour mettre à jour la base de données. Simple dans le principe, l'algorithmique de ce programme reste relativement complexe de part la multitude d'optimisations apportées afin de le rendre plus performant en terme de temps et qualité de réponse. Ce robot a trois tâches essentielles (UNJF/C2i, 2010) :

- (1) Trouver l'information sur le Web
- (2) Analyser cette information
- (3) Mettre à jour la base de données

2.4.2.2.1. Trouver l'information sur le web

- Le robot gère une liste de pages web (adresses URL) contenant toutes les pages web déjà connues dans sa base de données. Il interroge de manière cyclique cette liste de pages qui doivent être analysés ou ré-analysés. Une page est ré-analysée environ tous les 30 jours, mais cela dépend du moteur et d'autres éléments tels que la popularité de la page et sa fréquence de mise à jour.
- Lors de l'analyse d'une page, si le robot trouve un lien hypertexte vers une adresse URL qu'il ne connaît pas encore, il l'ajoute dans sa liste de pages web connues afin d'aller l'analyser ultérieurement.

2.4.2.2.2. Analyser l'information

Analyser une page consiste grossièrement pour le robot à aller récupérer cette page sur le web, puis à faire une indexation textuelle de son contenu. Nous parlons d'un robot, mais en réalité les moteurs de recherches font tourner en parallèle plusieurs robots. Les robots d'indexation de Google par exemple, aussi appelés Googlebots, peuvent analyser des milliers de pages simultanément.

L'indexation textuelle consiste à mémoriser dans la base de données pour chaque page des informations permettant de définir le contenu d'un document et de manière plus précise, de quoi parle ce document. En pratique, cela s'avère très compliqué parce que nous abordons des notions de sémantique d'un texte et qu'un programme ne peut pas comprendre réellement le sens d'un texte.

Choisir entre un annuaire et un moteur de recherche

Moteur ou annuaire, donc ? Aucune des deux approches n'est à exclure, car elles sont plutôt complémentaires et dépendent de ce que vous cherchez :

- **Un annuaire :** Ici, vous ne recherchez pas l'info, mais le site qui contient l'info.
 - Vous désirez identifier un site, une source générale d'information, un site ou une donnée officielle.
 - vous n'avez pas une idée précise et exhaustive de tout ce qui existe dans un domaine.
 - Vous voulez faire une recherche sur un domaine précis alors que les mots clés sont assez généraux. Dans ce cas, utilisez l'arborescence de l'annuaire pour descendre dans une catégorie générale correspondant à votre domaine de recherche, puis faites une recherche par mots clés localisée à une partie de l'arborescence.
- **Un moteur :**
 - Vous désirez identifier directement une information très précise (requête directement sur le contenu des pages).

2.4.2.2.3. Mettre à jour la base de données

Toutes les informations précédentes ainsi que la date de l'analyse sont enregistrées dans la base de données. Parallèlement à cette sauvegarde, certains moteurs de recherche, à l'instar de Google, conservent une copie de chaque page analysée sur des disques durs permettant ainsi aux utilisateurs d'accéder à la page telle qu'elle était au moment de l'analyse, si la page réelle venait à ne plus être accessible en ligne.

Pouvons donner quelques exemples de moteurs de recherche : Voila, Exalead, DeepIndex (français), Google, Bing, Msn, Yahoo et Altavista (anglais).

2.4.2.3. Les agents intelligents

2.4.2.3.1. La première génération : les méta-moteurs

Ils permettent d'interroger en une seule fois différents moteurs de recherche ou annuaires. Puis, pour faciliter la consultation, ils analysent les résultats, éliminent les doublons et peuvent vous proposer d'autres mots clés pertinents permettant de spécialiser votre recherche (UNJF/C2i, 2010).

Les méta-moteurs présentent quelques inconvénients :

- En général le problème à résoudre est plutôt celui de dénicher la bonne information au sein d'une grande quantité d'information, plutôt que de chercher dans encore plus d'information.
- Ces outils ne permettent pas de formuler de requêtes complexes tout simplement parce que les requêtes varient beaucoup d'un moteur à l'autre. Ces outils sont toutefois utiles lorsque l'objet de la recherche est extrêmement pointu et que la quantité d'information disponible sur le sujet est très limitée.

Exemples de méta-moteurs

- Metacrawler : L'un des méta-moteurs (moteur interrogeant plusieurs autres moteurs) les plus puissants du Web. A noter, l'intéressante fonction d'espionnage appelée MetaSpy qui permet de surveiller en temps réel les requêtes effectuées sur l'outil de recherche.
- Ixquick, Trouvez et Kartoo (français), Yippy, InfoSpace, PolyMeta et WebCrawler (anglais).

2.4.2.3.2. La seconde génération : des outils de veille

Il s'agit d'outils paramétrables qui font de la veille pour vous. Ces outils de veille se présentent soit sous forme de logiciels à installer sur son ordinateur, soit sous une forme directement accessibles via le web. Dans ce dernier cas vous êtes informés par email des nouveautés. Ils interrogent en général plusieurs moteurs à la fois.

Exemples d'outils de veille : Copernic et Googlealert.

2.5. Distinguer les difficultés

- Brouillage : la distinction entre support, contenu, mise en forme, graphisme.
- Fragmentation du contenu plus ou moins importante, selon les disciplines.
- Les rôles des acteurs et leurs compétences ne sont pas toujours identifiés ce qui influence sur la qualité du contenu surtout dans l'enrichissement des ressources Web.
- Information volumineuse, en évolution permanente.
- Ressemblance des ressources existantes dans le Web 'Plagiat'.

- Information gratuite et payante (tendance à plus d'information, plus rapide, moins chère, avec une frange d'information à valeur ajoutée payante).
- Identification des types des ressources (il se peut un type ou plusieurs dans le même besoin).
- Multilinguisme et couverture géographique mondiale.
- Le bruit (l'ensemble des documents pertinents non affichés lors d'une recherche).
- Le silence (l'ensemble des documents non-pertinents affichés suite à une recherche).
- Environnement non contrôlé contrairement à une bibliothèque ou une base de données où l'information est libre, mais non validée.
- Sélection de l'information parmi plusieurs sources : la responsabilité de la sélection des ressources ou des informations de qualité repose essentiellement sur l'utilisateur.
- Evaluation de l'information : être critique pour pouvoir évaluer le contenu intellectuel de l'information.

Pour répondre aux besoins des utilisateurs, il faut bien savoir ce qu'ils veulent chercher. Pour cela, le suivie d'une stratégie de recherche permet d'arriver aux résultats voulus.

Nous pouvons citer les étapes à suivre pour trouver des ressources pertinentes comme suit :

- 1- Choisir un outil de recherche
- 2- Type de la ressource (textes, images, données).
- 3- Le sujet recherché doit être précis.
- 4- Le contexte (universitaire, commercial, industriel...etc.).
- 5- L'objectif visé (faire un rapport, un cours, une étude, une enquête...etc.).
- 6- Chercher à identifier la ressource.
- 7- Ensuite, identifier soigneusement la ressource en s'assurant la nature du papier (éditeur, hébergeur, auteur, genre documentaire et sa nature, date de création et sa mise à jour).

Donc, il faut suivre une stratégie de recherche pour bien arriver aux besoins voulus, dans la section suivante, nous détaillons une stratégie de recherche standard pour pouvoir trouver les informations demandées.

2.6.Stratégie standard de recherche

Avant de commencer la stratégie de la recherche il faut bien connaître les différentes étapes de la recherche.

2.6.1. Les étapes de recherche

Selon (IRDES, 2012) les étapes de recherche se résument en :

- Préciser les objectifs de la recherche et bien formuler la question ;
- Choix des sources d'information pertinentes ;
- Définir les stratégies de recherche selon les sources interrogées ;
- Evaluer et sélectionner les sources d'information et les références obtenues ;
- Identifier et localiser l'information sélectionnée ;
- Hiérarchiser l'information et les documents collectés ;
- Présenter et communiquer l'information et les documents / Elaboration de la bibliographie finale (avec l'aide de logiciels bibliographiques : Reference manager, Endnote...)

Recommandations

- Le travail sur le vocabulaire est essentiel pour une recherche d'information efficace ;
- La quantité n'est pas synonyme de qualité ;
- S'interroger sur la qualité et la fiabilité des informations trouvées (sites internet validées...)

Dès que vous avez compris le fonctionnement des outils de recherche sur le Web, vous pouvez adopter une stratégie de recherche.

2.6.2. Quel type d'information recherchez-vous ?

Dans le cadre d'un travail de recherche universitaire vous pouvez avoir besoin de deux principaux types de recherche. Dans une recherche de type documentaire, il s'agit de (Rézeau, 2006) :

- retrouver les ouvrages d'un auteur dont vous connaissez le nom, ou bien un ouvrage précis, afin de l'emprunter à une bibliothèque, l'acheter, etc. ;
- retrouver tous les ouvrages et articles concernant un sujet.

La recherche documentaire permet de sélectionner l'information dans un fonds documentaire structuré en fonction de critères de recherche propres à l'utilisateur. Vous ferez ce type de recherche sur un catalogue de bibliothèque informatisé ou encore sur un site Web de type annuaire.

Dans une recherche de type "information", il s'agit plutôt de :

- retrouver des éléments d'information sur un thème donné : des statistiques, des données biographiques, une synthèse sur le sujet, etc. ;
- retrouver le texte d'une œuvre (source primaire) ou d'un ouvrage ou d'un article de critique (source secondaire).

Ce qui est visé dans ce deuxième type de recherche, c'est l'information plus que le document, le contenu plus que le contenant. Les moteurs de recherche et les annuaires du Web vous permettent de rechercher aussi bien des références de documents que les documents eux-mêmes ou encore des informations.

2.6.3. Définir les mots-clés

Ce sont les modes d'interrogation des bases de données bibliographiques qui exigent une recherche par mots-clés.

Pour trouver des documents sur un sujet dans une base de données, il faut utiliser les mêmes mots-clés que ceux utilisés par les indexeurs (IRDES, 2012). Les mots-clés sont regroupés dans un thesaurus, chaque thesaurus étant spécifique d'un domaine.

Etape 1 : Choix des mots-clés dans le thesaurus

Le thesaurus est une liste organisée de termes regroupés par domaines conceptuels et hiérarchisés en termes génériques et spécifiques. Les mots-clés sont définis lors de l'analyse du document. Ils décrivent le plus précisément possible le contenu du document. Nous bénéficions des thésaurus les caractéristiques suivantes : Suppression de la synonymie, Organisation des termes de façon hiérarchique, Association des termes de sens voisin et Normalisation des termes en précisant leurs sens.

L'utilisation d'un thesaurus permet de choisir un terme univoque pour être sûr d'obtenir le maximum de références par rapport à la question posée. L'orthographe doit être respectée.

Etape 2 : Equation de recherche et opérateurs booléens

C'est en recherchant et associant les bons mots-clés, que nous réalisons l'équation de recherche efficace. Les opérateurs booléens ET, OU, SAUF permettent de construire une équation de recherche, de croiser des mots-clés, d'en exclure...pour affiner la recherche et obtenir les références les plus pertinentes. ET Permet d'associer deux concepts ou plus, ainsi, permet de préciser la recherche, OU Permet d'élargir la recherche en sélectionnant les documents où apparaissent chaque terme ou plusieurs termes et SAUF Permet de délimiter la question pour réduire une interrogation en éliminant un des aspects de la question.

Recommandations

- Il est important de bien préparer la question à poser pour réduire les silences et les bruits.
- L'utilisation de la troncature ou astérisque permet de trouver les mots de même étymologie.

2.6.4. Formuler la requête

Lorsque vous aurez analysé votre demande d'information et établi une liste de mots-clés, il vous faudra rédiger votre équation de recherche en combinant ces mots-clés selon les règles expliquées précédemment dans la deuxième étape de la définition des mots clés (Rézeau, 2006).

2.6.5. Analyser les caractéristiques de son sujet de recherche

Il faut faire un bon choix du sujet de recherche (Rézeau, 2006) :

- comporte des mots ou des expressions ayant un sens bien caractéristique, précis et non ambigu ?
- ne contient pas de mots ou d'expressions précises ?
- concerne un point de vue d'ensemble sur un sujet relativement général ?
- concerne un aspect pointu d'un sujet plus vaste ou plus général ?
- peut être exprimé en des termes voisins ou synonymes ou comportant des variantes orthographiques ? comporte des termes polysémiques pouvant amener des résultats incongrus ?
- n'est pas bien clair dans votre tête ? Vous ne savez encore pas grand-chose sur le sujet ? Vous avez besoin d'être guidé ?

2.6.6. Choisir le bon outil de recherche

Voici quelques exemples de stratégie de choix d'un outil de recherche en fonction des caractéristiques de votre sujet de recherche analysées ci-dessus (Rézeau, 2006).

Le sujet de recherche doit respecter les caractéristiques suivantes :

- comporte des mots ou des expressions ayant un sens bien caractéristique, précis et non ambigu : mettez les expressions entre guillemets et faites une recherche dans un moteur comme Google. Pour les noms propres de personnes, essayez une requête avec prénom & nom à la suite puis une requête avec 'nom prénom'. La 2ème requête devrait réduire le nombre de résultats pour une personne très connue. Reformulez votre sujet en termes plus généraux, pensez en termes de domaine plus large et faites une recherche dans un annuaire comme Yahoo.
- ne contient pas de mots ou d'expressions précises : rajoutez des mots-clés ou des expressions entre guillemets pour restreindre le nombre de résultats dans les moteurs de recherche. Essayez de trouver des mots-clés précis dans les arborescences des annuaires ; utilisez les encyclopédies, les dictionnaires de synonymes pour établir une liste de mots-clés précis sur les thèmes concernant votre sujet de recherche.
- concerne un point de vue d'ensemble sur un sujet relativement général : n'utilisez pas un moteur de recherche mais un annuaire généraliste ou spécialisé 'thématique'.
- concerne un aspect pointu d'un sujet plus vaste ou plus général : utilisez les opérateurs booléens dans un moteur de recherche pour contrôler le trop-plein ou le manque de résultats.
- peut être exprimé en des termes voisins ou synonymes ou comportant des variantes orthographiques ou comporte des termes polysémiques pouvant amener des résultats incongrus : utilisez les opérateurs booléens dans un moteur de recherche.
- n'est pas bien clair dans votre tête ou vous ne savez encore pas grand chose sur le sujet ou vous avez besoin d'être guidé : allez fouiner du côté des annuaires généralistes ou spécialisés.

2.7. L'évaluation des résultats

Les documents traditionnels écrits (revues et ouvrages académiques) sont édités par des organismes (éditeurs, sociétés savantes) dont la notoriété apporte une garantie de sérieux du contenu de ces documents. Sur le Web, en revanche, n'importe qui peut mettre en ligne n'importe quoi. Avant d'utiliser (de citer) tout document trouvé sur le Web, il conviendra donc de prendre quelques précautions et de distinguer les documents en provenance de sites d'organismes connus ou reconnus d'une part (universités, sites de revues en ligne, sites de spécialistes reconnus dans leurs domaines) et ceux en provenance de pages personnelles mises en ligne par des individus (Rézeau, 2006).

Quel que soit le type de document trouvé et quel que soit le type du site qui l'héberge, il convient d'appliquer les règles habituelles de la critique externe du document, que nous le pouvons formuler au moyen des six (06) questions suivantes (Rézeau, 2006) :

Qui ? Quoi ? Où ? Quand ? Comment ? Pourquoi ?

Ces questions doivent être prises ensemble, et la décision de considérer comme fiables les informations doit venir de la convergence des réponses que nous obtenons.

➤ Qui est l'auteur du document ?

L'auteur d'une information est le premier responsable de celle-ci. Il est donc important de pouvoir l'identifier et déterminer ses compétences.

Questions

- L'auteur du document est-il clairement identifié : Dans le cas d'un article publié dans une revue en ligne, le nom de l'auteur (ou des auteurs) est clairement indiqué dans le sommaire des articles de la revue ainsi qu'en tête de l'article lui-même. Dans le cas d'un document faisant partie du site d'un organisme officiel ou commercial, le nom de l'auteur ou du rédacteur n'est pas toujours mentionné. Dans le cas d'un site personnel, le seul renseignement mentionné est parfois une adresse de courriel.
- Quelles sont les compétences de l'auteur ou des auteurs? Sont-elles clairement énoncées : Dans un article publié en ligne, l'appartenance de l'auteur à une institution, ses titres, son activité de recherche, sont normalement mentionnés.

- L'auteur fait-il référence à des publications antérieures, parues dans des périodiques imprimés ou publiés sur le Web et faisant autorité : Nous trouvons ces renseignements dans la rubrique *Références* des articles.
- Y a-t-il un moyen de vérifier le sérieux d'une organisation si elle n'est pas connue : Ce n'est pas facile, mais vous pouvez considérer que, si un organisme ou une source qui vous est inconnue est citée par plusieurs autres sources que vous savez être sérieuses, la source inconnue sera sérieuse elle aussi.
- Qui est l'éditeur du document : En dehors des sites personnels, l'éditeur d'un document ou d'un site Web est l'organisme ou le webmestre responsable du site. Le nom du serveur peut vous donner une indication sur la nature de l'organisme éditeur du site. Ce nom se trouve habituellement au début de l'adresse URL, entre `http://` et le premier `/`. Voici quelques exemples : `http://www.washingtonpost.com/` The Washington Post newspaper (site commercial), `http://www.harvard.edu/` Harvard University (site universitaire) et `http://dictionary.cambridge.org/` Cambridge University Press Dictionaries (éditeur).

➤ **Quoi ?**

La recherche d'une information à l'aide d'un moteur de recherche ou dans un annuaire peut amener un nombre élevé de résultats. Lors de la phase de "tri" de ces résultats vous devez parcourir les descriptifs et résumés du contenu des sites et documents trouvés en vous posant des questions vous permettant d'éliminer les résultats inintéressants pour votre travail de recherche. Nous posons les questions suivantes :

Pour le cas de la lecture des résumés : Quel est le sujet ou le thème général annoncé dans le résumé ? Ce sujet correspond-il bien à votre thème de recherche ?

Pour le cas d'un site Web : L'information fournie est-elle unique ? Existe-t-elle sur plusieurs sites? Existe-t-il des liens entre ces sites? A quel public s'adresse-t-on (spécialiste, général, initié, ...) ? Le traitement du sujet est-il exhaustif ?

➤ **Où ?**

En règle générale, l'adresse d'un site Web est intéressante, car elle apporte de précieux renseignements. Il faut en particulier examiner le nom du type de domaine, qui se situe entre `http://` et le premier slash `/`.

Par exemple, l'adresse <http://www.humanities.wisc.edu/contact/> appartient à un domaine de type "organisme d'éducation américain". Le Tableau 1 ci-dessous vous aidera à vous y repérer. Les fonctions avancées des moteurs de recherche vous permettent de spécifier un nom de domaine sur lequel effectuer votre recherche.

Les noms de domaines à trois lettres	
.com	domaine commercial (entreprises, grandes organisations commerciales, comme Microsoft.com, ...etc.)
.net	en principe domaine des réseaux (comme CERF.net, NSF.net) en réalité toutes sortes de sites
.org	organisations à but non lucratif (associations... mais aussi individus comme le site http://rezeau.org)
ces 3 premiers noms de domaines peuvent être hébergés n'importe où dans le monde	
Les codes nationaux à deux lettres	
.dz	Algérie
.be	Belgique
.fr	France
.eu	Europe
...etc.	...
sous-catégories ou sous-domaines	
.gov.dz	gouvernement algérien (ex. http://www.interieur.gov.dz)
.gouv.fr	gouvernement français (ex. http://www.education.gouv.fr/)
.gov.uk	gouvernement du Royaume-Uni (ex. http://www.royal.gov.uk/)

Tableau 1 Les adresses des sites Web

Pages personnelles

Nous aurions tort de penser que seuls sont valables les documents publiés sur les sites Web d'organismes officiels, d'universités ou d'éditeurs reconnus. De nombreuses pages personnelles contiennent des documents intéressants. En particulier de nombreux universitaires et chercheurs publient des documents sur leurs pages personnelles, que celles-ci soient hébergées sur le serveur de leurs universités ou sur un site complètement personnel et "non-officiel". Un indicateur courant d'une page personnelle publiée sur un site officiel est le tilde (~) comme par exemple dans l'adresse du site de Jack Lynch, professeur à Rutgers University : <http://andromeda.rutgers.edu/~jlynch/>.

➤ **Quand ?**

Un des avantages d'Internet est de fournir des informations presque en temps réel. L'inconvénient est le manque du recul nécessaire pour confronter l'information trouvée avec d'autres sources. Dans le domaine des lettres et de la littérature, les documents les plus récents ne sont pas nécessairement les plus pertinents.

En revanche, dans les domaines de la civilisation (contemporaine), de la linguistique et de la didactique, la date des documents trouvés ainsi que la date de dernière mise à jour des sites consultés sont importantes à prendre en considération.

Questions

- Quelle est la date de création du site? Quelle est celle de sa dernière mise à jour ? Cette date est-elle affichée dans la page ?
- Les liens vers d'autres sites sont-ils à jour et Existe-t-il beaucoup de liens brisés : Etant donné la labilité des sites sur le Web, les changements d'adresse, de nombreux liens ne mènent nulle part, nous disons qu'ils sont morts, ou brisés. Dans un site Web régulièrement mis à jour, le responsable du site vérifie périodiquement la validité des liens vers d'autres sites. Si vous constatez la présence de nombreux liens brisés sur un site, ce peut être un indice que le responsable du site n'est pas sérieux *ou* que le site n'est plus mis à jour.

➤ **Comment ?**

Les informations peuvent être rédigées ou chiffrées. Les informations chiffrées n'ont de valeur que si elles sont datées et si la source officielle est clairement indiquée. Des informations rédigées peuvent être présentées comme neutres ou polémiques.

A priori nous pouvons s'attendre à des informations plus objectives et neutres sur les sites d'organismes officiels (gouvernements, ministères, universités) que sur des pages personnelles. Cependant, toute information est susceptible d'être biaisée et il convient d'exercer votre esprit critique.

Bien que l'ergonomie et l'esthétique d'un document ou d'un site ne soient pas des indicateurs infaillibles de la validité de leur contenu, méfiez-vous des sites brouillons ou mal organisés : une présentation peu soignée est souvent la marque d'un contenu peu valable.

Questions

- Les ressources sont-elles bien organisées et présentées de façon logique ?
- S'il s'agit d'un site multimédia, quelle est la qualité visuelle et sonore des documents proposés?
- Les liens sont-ils complets, pertinents et appropriés : Les liens sont un peu comparables aux références dans les travaux scientifiques et universitaires. L'auteur d'un site ou d'un document Web qui vous invite à visiter d'autres sites ou documents semblables aux siens confirme la validité de ses affirmations en citant ses sources. Il vous invite en outre à comparer ses affirmations avec celles en provenance d'autres sources, ce qui peut être un gage de fiabilité.
- Sont-ils évalués ou commentés?
- Sont-ils relativement exhaustifs sur le sujet, ou servent-ils d'illustration?
- Ont-ils tous un rapport avec le sujet traité : Un certain nombre de pages personnelles ne donnent comme ressource rien d'autre que des listes de liens vers d'autres sites. Le plus souvent, le webmestre a seulement recopié ces listes sur d'autres sites semblables, sans se donner la peine de vérifier leur validité ou leur intérêt.
- Le type de langage employé est-il familier ou scientifique : Dans le cadre d'un travail de recherche universitaire vous ne pouvez pas utiliser (c'est-à-dire citer) des documents non scientifiques. Vous devez donc vous assurer que le document trouvé sur le Web offre toutes les garanties d'un article de type scientifique : article publié dans une revue en ligne, extrait d'ouvrage (thèse, etc.), extrait de cours d'un universitaire.
- Les documents sont-ils rédigés dans une langue correcte (orthographe et grammaire) : La correction de la langue d'un document trouvé sur le Web sera un élément déterminant quant à sa validité. Un texte rédigé dans une langue approximative et incorrecte ne peut évidemment pas être cité tel quel ; en outre, cette mauvaise qualité de la langue est très probablement l'indice d'une mauvaise qualité du contenu du document lui-même. Notez au passage que ce qui s'applique aux autres s'applique à vous-même. Quelles que soient les qualités de fond de votre travail de recherche, si sa rédaction est entachée de fautes de langue grossières votre lecteur aura du mal à croire à la valeur de votre argumentation.
- Le site propose-t-il un corpus cohérent et complet, ou bien est-il "en construction" : Un site peut être "en construction" et présenter néanmoins des documents tout à fait valables. Il faut juger les documents présents sur le site au cas par cas.

➤ Pourquoi ?

Nous s'intéressons ici aux motivations de l'auteur d'un site Web. Mettre des informations sur le Web coûte du temps et de l'énergie, et donc suppose que nous en retirons un certain profit.

En ce qui concerne les sites personnels, les motivations peuvent être : la satisfaction d'un ego qui sait qu'il sera lu, le plaisir altruiste de rendre service à la collectivité et la volonté prosélyte de diffuser des idées.

En revanche, les sites des organismes officiels ont une vocation naturelle à diffuser de l'information, et leurs présences sur le Web n'est que le prolongement naturel de leurs canaux d'information traditionnels (papier, médias...etc.). Il faudra être particulièrement méfiant envers les sites commerciaux et plus précisément à ceux qui n'annoncent pas clairement la couleur. Autant il est évident que les sites marchands de la FNAC, de Microsoft, de Amazon ou Alapage affichent clairement leurs buts de vous vendre leurs marchandises, autant de nombreux sites se présentent comme un service, une association, et déguisent habilement leurs intentions de vendre.

Questions

- Quel est le but du site ? Ce but est-il énoncé clairement ? Y a-t-il un résumé du contenu du site sur la page d'accueil : Recherchez sur la page d'accueil du site une rubrique intitulée *À propos de* ou *En savoir plus* (*About* en anglais).
- L'information est-elle gratuite, comme un service à la communauté? Y a-t-il des publicités sur le site? Si oui, ont-elles un rapport avec les informations que vous cherchez?
- Les publicités sont-elles clairement séparées du contenu des documents?
- Le site est-il clairement identifié comme site commercial (qui a un produit ou un service à vous vendre) ?

Remarques

En ce qui concerne les sites des revues scientifiques et universitaires en ligne, par exemple, il faut bien distinguer ceux qui proposent le texte complet des articles en ligne, gratuit ou payant, ou uniquement les résumés en ligne, les articles étant disponibles sous forme papier (abonnement payant à la revue).

2.8. Conclusion

L'Internet n'est qu'une source d'information parmi d'autres qui peut être très utile pour certains sujets, presque inutile pour d'autres. Lors d'une recherche documentaire exhaustive, il faut la coupler avec d'autres outils. C'est aussi une ressource qui évolue vite et qu'il faut suivre attentivement sous peine d'y perdre ses repères

Face à l'énormité exponentielle du web devenu la première source d'information mondiale, l'internaute qui désire obtenir la réponse à une question peut recourir à différents outils ; nous distinguons entre annuaires, moteurs de recherche, méta-moteurs et agents ' intelligents '. Les plus utilisés sont les moteurs de recherche et en particulier Google et Yahoo.

Notre problématique de recherche se focalise sur cette perspective, c'est-à-dire travailler avec un moteur de recherche en réalisant un module qui va chercher les documents les plus pertinents et les télécharger dans un support local. Premier problème rencontré, les résultats trouvés peuvent n'avoir aucune relation avec la requête saisie. Pour cela, il faut réaliser un autre module qui va s'occuper du filtrage des documents téléchargés en se basant sur la notion du Web sémantique et plus précisément sur une ontologie de domaine. Nous avons rajoutés aussi une nouvelle notion nommée règle sémantique où chacune d'elle est définie dans l'ontologie de domaine. Chaque concept de l'ontologie peut avoir plusieurs règles sémantiques.

L'objectif du projet PrOALDoF-Web est de s'adapter à la dynamique intrinsèque du Web en termes d'informations et de contenu éducatif, d'usages (plates-formes, environnements, utilisateurs). Au sein de ce projet, nous proposons un outil Web avec les caractéristiques suivantes:

- Recherche Web intelligente et filtrage décisif, basée sur les ontologies et les règles sémantiques afin de retrouver et extraire les sous fragments des documents les plus pertinents ;
- Adaptation automatique des sous fragments des documents en fonction des contraintes des utilisateurs ' profils des apprenants '.

Le chapitre suivant se focalise sur les systèmes éducatifs en ligne dans lequel nous avons appliqué notre approche de recherche.

3. Les systèmes éducatifs en ligne

3.1.Introduction

Un peu partout dans la littérature informatique, il est de plus en plus question du E-Learning, de la formation en ligne. D'où vient cet intérêt grandissant pour ce type d'apprentissage? (Haut, 2003).

Les Technologies de l'Information et de la Communication (TIC) sont de plus en plus utilisées et de façon variable par les établissements d'enseignement supérieur à travers le monde. Elles sont intégrées autant dans un enseignement présentiel que dans un enseignement ouvert et à distance.

Ces TIC induisent des changements importants dans les pratiques pédagogiques, en transformant les modalités de communication, de travail et de formation (Gharbi et Ouerfelli, 2008). Dans ce fil d'idées, plusieurs auteurs (Karsenti, Savoie-Zajac et Larose, 2001) et (Charlier et Peraya, 2003) affirment que l'intégration des TIC dans l'enseignement implique des modifications des pratiques des enseignants et touche à leurs représentations de l'apprentissage et à leurs modalités de collaboration et d'évaluation.

L'usage des TIC a permis de mettre en place un nouveau mode d'enseignement, celui de la formation à distance également appelée *E-formation*, *E-Learning* ou formation en ligne. Ce type de formation peut se définir par rapport à la formation présentielle, comme étant une formation qui se caractérise par la délocalisation spatio-temporelle. Les apprenants ne sont pas obligés d'être présents dans les mêmes lieux et en même temps que les enseignants (Gharbi et Ouerfelli, 2008). La formation à distance est renforcée par l'avènement des réseaux et des technologies fondées sur Internet (outils de communication synchrone et asynchrone, Web, etc.).

A partir d'un centre ou une université, nous pouvons rejoindre plusieurs universités dans leur quête de briser les contraintes spatio-temporelles de l'enseignement universitaire traditionnel et de profiter de l'expansion fulgurant des TIC qui permettent de concevoir et d'acheminer des cours interactifs sur le web. La mise en place et l'expérimentation d'une plate-forme d'enseignement à distance nécessite une première évaluation des usages et des attentes des enseignants (Gharbi et Ouerfelli, 2008).

L'évaluation du ou des systèmes éducatifs doit obligatoirement se déployer dans plusieurs dimensions (Haut, 2003) : évaluations des enseignements, évaluation des politiques éducatives, évaluation de la gestion, évaluations bilans nationales, comparaisons internationales, statistiques...

Mais, nous avons essayé de définir dans un premier temps l'enseignement à distance, exposer son évolution historique et étaler ses principales caractéristiques. Dans la partie suivante, nous discutons les caractéristiques générales d'un E-Learning en détaillant les principes de base de ce concept. Nous plongeons ensuite au cœur du E-Learning en décrivant les différents éléments qui le composent, les acteurs qui l'animent. Enfin, les outils qui permettent de gérer ces formations.

3.2.Définition

L'UNESCO définit l'enseignement à distance comme un 'mode d'enseignement, dispensé par une institution, qui n'implique pas la présence physique du maître chargé de le donner à l'endroit où il est reçu, ou dans lequel le maître n'est présent qu'à certains moments ou pour des tâches spécifique'.

Pour sa part, l'AFNOR définit cet enseignement comme étant un mode de formation à distance 'conçu pour permettre à des individus de se former sans se déplacer sur le lieu de la formation et sans la présence physique d'un formateur. La formation à distance est incluse dans le concept plus général de formation ouverte et à distance'.

Geneviève Lameul (Riffaut et Rostaing, 2002) note l'apport des outils de communication dans un processus de formation à distance. Elle considère que l'utilisation des différents moyens de communication permet de 'rompre avec les trois unités de temps, lieu et action, en permettant aux apprenants de se former sans avoir à se déplacer dans un lieu spécifiquement identifié pour la formation, en entrant en contact avec le formateur-accompagnateur ou l'enseignant par l'intermédiaire des moyens de communication'.

Dans ce travail nous retenons le concept de l'enseignement à distance ou de formation à distance comme un dispositif composé d'un ensemble de ressources matérielles, techniques, humaines et pédagogiques, mis en place pour fournir un enseignement à des individus qui sont distants.

Nous rejoignons ainsi, Daniel Peraya (Peraya, 2005) qui considère que 'La formation à distance parce qu'elle dissocie dans le temps et dans l'espace le processus d'enseignement/apprentissage, apparaît d'emblée comme une formation en différé et, en conséquence, elle doit nécessairement se concevoir et être mise en œuvre comme une formation médiatisée. Autrement dit, toute formation à distance recourt nécessairement à des artefacts techniques, à des dispositifs médiatiques anciens ou nouveaux, analogiques ou numériques : livres ou manuels imprimés et illustrés, radiodiffusion, enregistrements sonores sur cassettes, émissions de télévision en direct ou magnétoscopées, technologies de l'information et de la communication. C'est ainsi que seront mis à disposition des apprenants les ressources pédagogiques, les différents outils de communication et de collaboration, les activités d'apprentissage et les moyens de les réaliser. C'est à travers de tels dispositifs que les apprenants communiqueront et collaboreront, qu'ils se verront aider dans leurs travaux'. Ce mode d'enseignement requiert donc des techniques spéciales de formation, de conception de cours et des moyens de communication.

Brandon-Hall (Brandon-Hall, 2013), un spécialiste de l'E-Learning, en donne la définition suivante: 'Le E-Learning est un programme de formation qui est accessible via un navigateur Internet, à travers Internet ou un Intranet. En utilisant le Web (ou le web d'un Intranet) pour la formation, nous faisons fait référence par définition à l'environnement visuel et interactif propre à Internet (Haut, 2003)'.

Cette définition, communément acceptée, est cependant trop réductrice en ce sens qu'elle se limite à la technologie utilisée dans l'E-Learning et ne tient pas compte de ses aspects conceptuels qui nous paraissent être l'essence même de ce nouveau mode de formation.

3.3.Aperçu historique

Historiquement, l'enseignement à distance a développé un mode de formation axé sur l'imprimé dans un premier temps, ensuite sur des technologies de mass média avec la radio et la télévision et enfin sur l'informatique, les réseaux et les services numériques interactifs *on et off line*. Dans ce même ordre d'idées, Nipper cité par Peraya (Peraya, 2005) propose trois repères chronologiques dans l'histoire de la formation à distance qu'on rappelle ici brièvement : L'imprimé qui marque le début de la formation à distance et constitue la base des cours par correspondance.

Selon Peraya (Peraya, 2005) le dispositif se fondait sur les technologies du livre et de l'imprimé : le texte écrit et ses différents para-textes (photographies, schémas, dessins, graphiques, cartes, etc.). Dès les années 60, s'ouvre l'ère du multimédia caractérisée par un usage de différents médias (imprimé, radio, télévision, vidéo) complémentaires et coordonnés en vue d'un objectif pédagogique commun. Bien que les émissions éducatives soient utilisées aux Etats-Unis d'Amérique depuis la première moitié du XXème siècle (environ 1920 pour la radio et 1950 pour la télévision), ce ne sont devenus des composants habituels des cours d'éducation à distance qu'à partir de 1970, lorsque les télé-distributeurs et les câbles de télévision se sont généralisés. A ce moment, l'audioconférence devenait de plus en plus populaire auprès de nombreuses universités comme l'affirment Charlier et al (Gharbi et Ouerfelli, 2008). A la fin des années 1980, les progrès de la micro-informatique et des télécommunications ont permis à l'enseignement à distance d'effectuer une nouvelle mutation.

En effet, deux nouvelles formes de technologies sont apparues : le multimédia et les télécommunications. Le développement des télécommunications permet à l'apprenant de dialoguer avec l'enseignant et les autres apprenants par visioconférence, par courrier électronique, via des forums, d'exploiter les ressources pédagogiques du web, de s'auto évaluer en ligne... Ainsi, avec le développement des télécommunications et des réseaux permettent le commencement de l'époque contemporaine, celle des plates-formes d'enseignement à distance et du campus virtuel défini (Peraya, 2005) comme un environnement unique intégrant différentes fonctions d'information, de communication (synchrone ou asynchrone), de collaboration, de gestion et d'apprentissage.

3.4. Les principes du E-Learning

Quatre principes fondamentaux forment la base du concept de E-Learning (Haut, 2003) :

3.4.1. Premier principe: utilisation des nouvelles technologies

Le premier principe porte sur l'utilisation des Nouvelles Technologies de l'Information et de la Communication (NTIC), interactives et accessibles à un très large public. Ces nouvelles techniques permettent : **Premièrement**, de s'affranchir des limites de lieu et de temps. Grâce à l'Internet (ou à un Intranet), il n'est plus nécessaire que l'ensemble des élèves soient réunis dans un même lieu avec le professeur.

Dans ce sens, les distances sont abolies et, de manière analogue, le temps est différent puisque chaque élève peut, dans une certaine mesure, gérer lui-même son rythme d'apprentissage et adapter son planning en toute liberté. Il est cependant conseillé de prévoir à intervalles réguliers dans le temps des 'points de rencontre' virtuels ou non afin de permettre au professeur de corriger des dérives ou des erreurs éventuellement constatées en cours d'apprentissage.

Et **deuxièmement**, d'offrir plusieurs moyens de communication entre l'apprenant et l'enseignant, de même qu'entre les apprenants : des forums de discussion peuvent, par exemple, être créés pour favoriser les échanges d'idées entre les participants. Ces forums peuvent d'ailleurs perdurer au delà du cours pour permettre aux apprenants de continuer à partager leurs expériences. Et donc, de profiter de toutes les possibilités du multimédia pour proposer un enseignement enrichi.

On sait depuis longtemps que l'association de textes et d'images renforce le processus d'apprentissage et de mémorisation en provoquant un double encodage dans la mémoire. Aujourd'hui, le contenu des cours peut encore être enrichi en y intégrant des extraits sonores, des vidéos, des animations, des simulations et en favorisant l'interaction avec l'apprenant. Toutes ces techniques permettent d'accroître l'intérêt et l'implication de l'apprenant avec comme objectif une meilleure mémorisation de la matière et d'élargir l'espace d'apprentissage en proposant des compléments à la formation sous forme de liens vers des sites Internet où l'apprenant pourra trouver de la documentation plus poussée sur certaines parties du cours qui l'intéressent plus particulièrement. Dans le même ordre d'idée, des liens vers une base de connaissances interne à l'entreprise peuvent également être proposés aux apprenants.

3.4.2. Deuxième principe : organisation du contenu

Le second principe concerne l'organisation du contenu pédagogique. Ce contenu est découpé en modules de telle sorte que l'on puisse construire des parcours d'apprentissage adaptés aux élèves en fonction de leurs connaissances de base, de leurs expériences, de leurs objectifs ...etc.

Cette personnalisation de la formation, basée sur la pédagogie de la construction, est probablement la caractéristique la plus importante de l'E-Learning mais aussi celle qui est la plus difficile à mettre en œuvre.

Sur la base d'une analyse globale du contenu du cours, il s'agit de définir des sous-ensembles ou modules cohérents et de définir une progression logique (un chemin d'apprentissage) d'un module à l'autre de telle sorte que les conditions suivantes soient satisfaites : Les pré-requis d'un module doivent être contenus dans le(s) module(s) précédent(s) (sauf les connaissances de base bien sûr). Chaque module apporte à l'apprenant un ensemble de connaissances qui viennent de s'ajouter, de manière cohérente, aux connaissances acquises dans les modules précédents; cela signifie que si l'apprenant arrête le cours au terme d'un module donné, l'enseignement reçu serait certes incomplet mais formerait un tout utile. Il est d'ailleurs souhaitable que, dans la mesure du possible, les modules soient conçus de telle façon qu'après avoir terminé un module avec succès, l'élève puisse mettre en pratique, dans son travail quotidien, la connaissance acquise. L'assimilation du contenu de chaque module doit pouvoir être évaluée par des tests en fin de module. Ces tests seront proposés à l'apprenant sous forme de questions/réponses, de questionnaires à choix multiples, de problèmes à résoudre, de simulations, d'études de cas...etc. Certains modules sont, dans leurs contenus, redondants par rapport à d'autres modules tout en présentant ce contenu différemment : l'utilité de ces modules est de permettre à l'apprenant de revoir, sous une autre forme pédagogique, un contenu qu'il n'a pas bien assimilé (ce qui sera mis en évidence par les tests de fin de module).

3.4.3. Troisième principe: le rôle de l'enseignant

Dans le E-Learning, le rôle du professeur est radicalement différent de celui qu'il joue dans un enseignement traditionnel : il n'y a plus d'exposé ni de structure fixe de cours, l'enseignant y joue un rôle de tuteur (ou coach), il doit plus se comporter en professeur particulier attentif à la progression de chaque apprenant. Le tuteur, dégagé des tâches traditionnelles, pourra consacrer la totalité de son temps à suivre, pratiquement en temps réel, l'évolution de chaque apprenant. La technologie sous-jacente lui donne en effet la possibilité de connaître à quel point du parcours pédagogique se trouve chaque élève, d'identifier les parties de la matière qui posent des problèmes à l'un ou à l'autre, de prendre connaissance des questions qui lui sont adressées, de consulter les forums de discussion relatifs au cours dispensé, ...etc.

En fonction de tous ces éléments qu'il peut récolter, le tuteur aura pour tâche de conseiller les participants au cours, de moduler leurs parcours pédagogiques en modifiant éventuellement l'enchaînement des différents modules, de susciter des discussions au sein d'un forum de discussion et même d'organiser sporadiquement des rencontres virtuelles ou réelles.

Comme nous pouvons constater, la fonction de tuteur requiert non seulement des compétences pédagogiques mais aussi des compétences techniques et des capacités de dialogues orientées vers l'aide à apporter aux élèves.

3.4.4. Quatrième principe: la collaboration entre apprenants

Nous considérons aujourd'hui comme acquis que l'apprentissage en groupe donne généralement de meilleurs résultats que l'apprentissage isolé.

Par rapport à une formation classique, les élèves d'un cours dispensé par E-Learning ne sont plus physiquement rassemblés dans un même local. Comme il est important de ne pas perdre la valeur ajoutée que représente la communication formelle et informelle entre les apprenants, les plates-formes de E-Learning fournissent des outils qui permettent de créer des classes virtuelles où des groupes d'apprenants vont pouvoir 'se rencontrer', échanger des idées, confronter leurs expériences.

Le Tableau 2 ci-dessous (Haut, 2003) résume les principes de base du E-Learning en comparant succinctement une formation traditionnelle et une formation par E-Learning selon certains critères.

Formation traditionnelle	E-learning
Organisation du temps	
Les modalités de formation sont fixées d'avance et imposées à tous avec une rigidité qui ne tient pas toujours compte de la réalité..	L'apprenant organise <u>lui-même</u> sa formation <u>pendant son temps de travail</u> .
Organisation du parcours pédagogique	
Progression linéaire et prédéfinie.	Progression <u>adaptée à chacun</u> en fonction de la situation.
Connexion avec les activités opérationnelles	
Sans connexion immédiate.	<u>Intégrée</u> aux activités opérationnelles.
Modalités	
Activité ponctuelle, ayant lieu à un moment précis et à un moment imposé.	Activité <u>continue</u> .
Contenu	
Le contenu est le même pour tous et organisé pour satisfaire le plus grand nombre. Il ne tient compte des changements dans l'entreprise qu'avec retard.	Le contenu tient compte <u>des besoins individuels</u> et des interactions entre participants et tuteur. Le contenu <u>peut s'adapter plus rapidement</u> à l'évolution des besoins dans l'entreprise.

Tableau 2 Comparaison entre une formation traditionnelle et le E-Learning

3.5. Les acteurs et leurs rôles dans le E-Learning

Les différents acteurs d'un cours E-Learning peuvent être classés par catégories suivant le rôle qu'ils seront amenés à jouer. Nous distinguerons les intervenants suivants (Haut, 2003):

- L'**apprenant** est la personne qui suit la formation proposée. Il organise lui-même son parcours d'apprentissage en gérant la façon dont il y consacre son temps. Tout au long de cet apprentissage, il a la possibilité d'interagir avec les autres apprenants et avec le tuteur du cours soit pour échanger son expérience avec les autres soit pour obtenir un aide parfois nécessaire.
- Le **tuteur** joue un rôle fondamental qu'on a précédemment expliqué. Pour pouvoir mener à bien ses différentes tâches de suivi des élèves, de conseil, de modérateur de discussion..., il est évident que le tuteur doit être expert dans la matière enseignée.
- Le rôle d'**évaluateur** est souvent tenu par le tuteur lui-même; il consiste à corriger les tests et les exercices proposés et à attribuer des notes aux élèves. Ces notes serviront à moduler le parcours d'apprentissage des apprenants et pourront également être exploitées ultérieurement pour identifier les parties du cours qui posent problème afin d'améliorer constamment ce dernier.
- L'**auteur** ou **concepteur** du cours a la lourde tâche non seulement de créer le contenu du cours mais également de construire l'architecture de ce contenu : définition des différentes filières, définition des modules et structuration du cours en déterminant l'enchaînement de ces modules pour chaque filière.
- Le **réalisateur** du cours prend en charge la traduction informatique de tous les éléments du cours : développement des interfaces, présentation du contenu du cours, système de navigation, construction des tests et automatisation de leur évaluation, réalisation des animations et des simulations, mise à disposition d'outils de communication (messagerie, forums, chat, ...etc.). Cette réalisation se fera souvent en équipe et en lien étroit avec l'auteur du cours et le tuteur. Cette fonction requiert des compétences techniques étendues comprenant la maîtrise des outils techniques que nous détaillons plus loin. Et en plus de ces compétences techniques, il est utile de disposer, au sein de l'équipe, d'un collaborateur avec une expérience en infographie et en construction de sites Web.

- Le **support technique** doit gérer l'infrastructure matérielle, assurer la disponibilité du cours et aider les autres intervenants dans les problèmes techniques qu'ils peuvent rencontrer. C'est lui également qui assure la gestion des bases de données liées au cours.
- Le **gestionnaire** du (ou des) cours assume notamment la gestion des plans de formations et des inscriptions, l'enrôlement et le suivi global des élèves.

Certains de ses rôles pourront être assurés par une même personne et inversement, il sera parfois nécessaire d'avoir plusieurs personnes pour assumer une même fonction.

3.6. Les outils dans un E-Learning

Chaque catégorie d'intervenants doit pouvoir disposer d'outils qui l'aident à assumer sa tâche. L'apprenant n'a en principe besoin que d'un PC raisonnablement puissant, d'une connexion Internet ou Intranet et d'un navigateur Internet pour accéder au cours.

Il est de bonne pratique de vérifier, tout au début du cours, qu'il dispose, sur son PC, de tous les plugins qui seront nécessaires pendant le cours et de mettre à sa disposition ceux qui lui manquent.

Pour les besoins de certains cours ou pour pouvoir réaliser certains exercices, l'élève devra parfois télécharger du site de cours des outils particuliers à une discipline. Dans ce cas, nous devons veiller à ce que ces outils soient aisément accessibles et que leur installation soit bien documentée.

Dans une plate-forme typique de E-Learning, nous trouvons des outils de management des formations (LMS 'Learning Management System'), des outils de création des modules de cours (LCMS 'Learning Content Management System') et des outils destinés aux différentes formes de communication.

La Figure 3 en page 46, (Hyperwave, 2013) illustre schématiquement les différents modules d'une plate-forme classique. Il est à noter cependant qu'il n'y a pas toujours une frontière bien nette entre les différents modules et qu'il n'est pas non plus nécessaire d'acquérir une plate-forme complète pour se lancer dans une expérience de E-Learning. En particulier, nous trouvons des outils de création de contenu (Authoring) qui offrent déjà des possibilités de création de modules de cours (Haut, 2003).

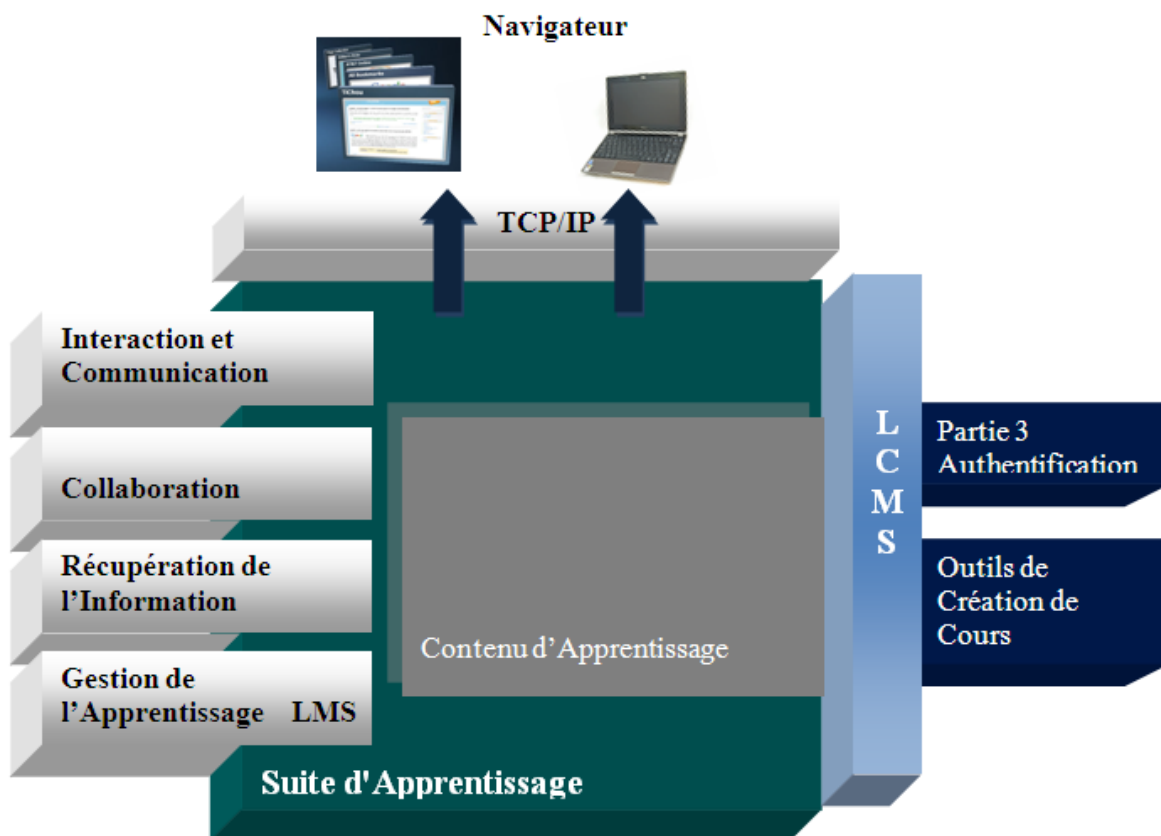


Figure 3 Les différents modules d'une plate-forme classique de E-Learning

3.6.1. Les outils de gestion des cours (LMS)

Les LMS se réfèrent à l'ensemble des outils informatiques mis à disposition des gestionnaires pour administrer les formations. Ces outils offrent généralement les fonctionnalités suivantes :

- Gestion des catalogues de formations, regroupement de ces formations par thèmes, tenue de fiches descriptives.
- Planification des cours notamment pour assurer la disponibilité d'un tuteur tout au long de la période pendant laquelle le cours est accessible.
- Gestion des inscriptions des élèves, y compris, éventuellement, les procédures de facturation des cours.
- Suivi du parcours d'apprentissage des élèves : mémorisation des cours suivis, enregistrement des résultats obtenus lors des différents tests, analyse de ces résultats dans l'objectif de proposer le parcours le plus adéquat en fonction des objectifs de l'apprenant.

- Génération de rapports particuliers destinés à l'élève et de rapports globaux à l'usage des tuteurs et des gestionnaires.
- Délivrance de diplômes éventuellement reconnus par une autorité.
- Tenue de statistiques diverses : fréquentation des cours, durées nécessaires à l'assimilation des différents modules, niveaux de réussite ou estimation de difficulté des tests, pertinence des critères de passage d'un module à l'autre...etc.
- Aide à la gestion et au monitoring de l'infrastructure informatique de façon à garantir l'accessibilité des ressources techniques nécessaires au suivi des cours (serveurs, réseau, outils de communication et de collaboration...etc.).
- Gestion des ressources pédagogiques : disponibilité des contenus de cours, des documents proposés en compléments d'information, vérification régulière des liens vers des ressources externes ou vers des bases de connaissances.

3.6.2. Les outils de création des cours (LCMS)

La distinction entre les outils de création de cours et les outils de création de contenu n'est pas toujours très claire et suivant les plates-formes, ces deux outils sont plus ou moins regroupés. Pour fixer les idées, nous pouvons définir les outils de création de cours comme ceux qui offrent les fonctionnalités nécessaires à l'assemblage en modules des pages de contenu (comme l'on assemblerait en chapitres les pages d'un livre) et à la création du cours à partir de ces modules. Nous y trouverons donc :

- Facilité d'assemblage des pages en modules avec construction automatique de la navigation à l'intérieur du module. Des modèles de présentation du contenu qui aideront à assurer une présentation uniforme à travers tout le cours. Des guides interactifs 'wizards' aident généralement à la construction d'une présentation personnalisée.
- Un aide à la création des différents types de tests (questionnaires à choix multiples, tests ouverts, exercices...), à la présentation des résultats de ces tests et à leurs corrections. Ce sont ces mêmes outils qui assureront le stockage des résultats dans la base de données 'ad-hoc' du LMS.
- Des outils permettant de construire plus ou moins facilement la logique de navigation d'un module à l'autre en fonction de la filière de cours et des résultats aux tests de fin de module.

Un des apports du E-Learning par rapport aux cours traditionnels est de pouvoir stimuler divers canaux d'apprentissage pour optimiser la mémorisation et l'apprentissage chez l'élève. Il est en effet de peu d'intérêt d'offrir un cours de E-Learning où nous aurions simplement regroupé en modules et formaté en HTML 'HyperText Markup Language' les pages d'un livre de cours ou les diapositives d'une présentation PowerPoint.

Les technologies actuelles permettent d'enrichir de manière importante le contenu des pages de cours ou des 'grains pédagogiques'. C'est à ce niveau qu'interviennent les outils de création de contenu proprement dit (que l'on qualifie généralement d'outils d'authoring). Ils permettent de construire les images, les vidéos et les bandes-sons que nous voulons les associer au texte. Ce sont également eux qui sont utilisés pour construire les interactions avec l'élève et pour développer des simulations, les composantes majeures (et une partie importante de la valeur ajoutée) d'un cours E-Learning attractif. Dans la majorité des plates-formes d'E-Learning, des outils tiers sont utilisés à cet effet.

3.6.3. Les outils de communication et de collaboration

Une plate-forme de E-Learning doit permettre les échanges entre les apprenants, le travail en commun, les liens avec les tuteurs, les accès à des bases de données de connaissances. Les outils de communication asynchrone comprendront un système de messagerie et un forum de discussion. Il est préférable que le système de messagerie soit dissocié de la messagerie habituelle des participants et soit intégré à la plate-forme d'E-Learning. De cette façon, il sera plus simple pour l'apprenant de gérer distinctement les échanges de courrier relatifs au cours qu'il suit. Les forums de discussion seront limités à une communauté bien définie de façon à éviter de 'polluer' les forums par des sujets de peu d'intérêt. Typiquement, une telle communauté comprendra tous les participants à une filière donnée d'un cours particulier. A côté des outils asynchrones, nous trouverons généralement deux autres outils de communication synchrone : (1) Les logiciels de 'chat' (de 'bavardage') qui permettront aux intervenants de discuter entre eux en temps réel. (2) Les logiciels de 'classe virtuelle' (ou de 'tableau blanc') : avec ce type de logiciel, tous les apprenants voient, sur l'écran de leurs ordinateurs, le contenu affiché sur celui de l'enseignant (le tuteur). L'enseignant peut ainsi présenter de nouveaux documents, les commenter et les annoter; il peut également 'passer la main' à un des participants, lui donnant de cette façon la possibilité de partager ses idées. Enfin, la possibilité de travailler en mode déconnecté, après téléchargement du contenu des modules nécessaires, est une fonctionnalité intéressante offerte par certaines plates-formes.

3.7. Conclusion

Idéalement, le E-Learning est une méthode de formation basée sur la pédagogie dite 'de la construction' où les technologies les plus récentes sont utilisées pour construire un contenu pédagogique propre à stimuler tous les canaux de mémorisation, et pour permettre à chaque apprenant de choisir librement un parcours d'étude qui correspond le mieux à ses capacités et à ses attentes, tout en étant guidé et en ayant, en permanence, la possibilité de faire appel à un tuteur et de partager son expérience avec la communauté des autres apprenants (Haut, 2003). Après avoir détaillé les principes qui caractérisent le E-Learning, son architecture, les éléments techniques qui le constituent et les acteurs impliqués dans un E-Learning ainsi que leurs rôles, le E-Learning a plusieurs avantages par rapport à une formation traditionnelle, non seulement a propos des coûts qui peuvent être diminués d'une valeur importante (pour autant qu'il y ait un nombre suffisant d'apprenants) mais également au niveau de la qualité, notamment par la richesse potentielle du contenu et la personnalisation de l'enseignement.

Le E-Learning ne remplacera pas l'enseignement traditionnel mais nous pensons qu'il pourrait avantageusement le compléter sinon le remplacer dans certains domaines où il est fondamental d'allier des connaissances et des expériences. Il faudra en tout cas, à l'avenir, tenir compte des avantages considérables que les nouvelles technologies apportent au monde de la formation.

Le développement de la formation en ligne conduit les apprenants à avoir moins de contacts avec les enseignants et les amène à être plus autonomes, à être davantage acteurs de leur formation. Parallèlement, ils peuvent disposer de plus en plus de documents ou de ressources, que celles-ci aient été produites dans le cadre de la formation qui leurs est destinée ou qu'elles soient directement disponibles sur le web. Il est cependant souvent difficile, pour les apprenants et même pour les concepteurs de formations à distance, d'identifier les ressources pertinentes et de les organiser dans des ensembles cohérents. Cela a conduit à des initiatives visant à constituer des banques de ressources pédagogiques ou des 'Repositories' qui sont gérées et exploitées par des réseaux de contributeurs.

Pour que l'apprenant puisse utiliser l'ensemble des ressources mises à sa disposition avec efficacité, il faut satisfaire deux exigences a priori antinomiques, à savoir celle qui vise à lui accorder une certaine liberté de choix, et celle qui lui évite de se disperser a l'égard de ses connaissances du moment.

Les banques de ressources sont plutôt destinées aux concepteurs d'activités pédagogiques ou aux enseignants pour qu'ils mettent en place rapidement des formations adaptées. Notre problématique de recherche se place dans cette perspective, en sélectionnant les documents pertinents à partir du web et en filtrant ces derniers afin de construire un domaine d'enseignement, et aider les concepteurs ou les enseignants à trouver rapidement les ressources pertinentes.

Une application E-Learning est mise en ligne via l'utilisation du Web. De ce fait, elle partage le même problème de pertinence avec le Web. Face à un géant nombre d'internaute (2,405,518,376 jusqu'à 30 juin 2012) (IWS, 2012), l'internet est devenu un moyen de communication et de recherche quotidienne très important, mais avec plus d'un 1000 milliards de pages sur le web que Google a répertorié (Duffez, 2008), nous sommes devant une obligation pour changer la méthode d'indexation, pour cela les ingénieurs qui font les robots permettant l'indexation des pages Web dans les moteurs de recherche changent ses principes, en basant sur le Web sémantique.

Dans le chapitre suivant, nous détaillons les liens qui unissent les applications de E-Learning et le Web Sémantique, ainsi nous présentons le potentiel de l'ontologie et de l'ingénierie ontologique pour augmenter l'intelligence et la communication dans les applications E-Learning et surtout pour faciliter la recherche des informations les plus pertinentes et aider les constructeurs de cours à filtrer les ressources pédagogiques.

4. E-Learning et le Web Sémantique 'les ontologies'

4.1. Introduction

Une application E-Learning est mise en ligne via l'utilisation du Web. Compte tenu de la diversité et la croissance exponentielle des ressources pédagogiques utilisées dans le cadre d'une formation de type E-Learning, il est de plus en plus difficile de trouver les documents pédagogiques pertinents. Une application E-Learning partage donc le même problème de pertinence avec le Web lorsque les apprenants veulent accéder au savoir mis à leur disposition. Concevoir et développer des outils spécifiques pour faciliter l'accès aux documents pédagogiques et leurs intégrations devient une nécessité. Parmi les efforts visant à résoudre le problème de pertinence nous trouvons la notion de 'Web Sémantique'. Cette nouvelle génération du Web, apparaît comme une technologie prometteuse pour implémenter le E-Learning. Le Web Sémantique constitue un environnement dans lequel les agents humains et machine vont communiquer selon une base sémantique (Berners-Lee, 2000). Une des caractéristiques principales est la compréhension partagée basée sur un squelette d'ontologie. L'ontologie permet l'organisation du matériel d'apprentissage autour de petites pièces d'objets d'apprentissage sémantiquement annotés (enrichis). Les items peuvent être facilement organisés en des cours d'apprentissage (rapide et juste à temps) et livrés à la demande à l'utilisateur selon son profil et les besoins d'affaire (pertinence) (Boutemedjet, 2004). L'utilisation des ontologies est une piste privilégiée pour l'analyse, la conception et le développement des applications E-Learning.

Notre problème de recherche vise à construire une ontologie de domaine contenant les concepts avec des annotations bien précises, ces concepts vont être le moyen de comparaison dans les documents en premier lieu, ensuite une deuxième comparaison avec les annotations, afin d'extraire les parties pertinentes de chaque document, et en conclusion les bonnes parties nécessaires (les plus pertinentes) vont faire la construction du cours voulu. Puis, nous avons une ontologie pédagogique qui sert à adapter le cours construit aux apprenants.

Le présent chapitre définit le web sémantique et les ontologies, puis il décrit les objectifs de la nouvelle génération du Web, son architecture globale, ses langages candidats, à savoir : XML, RDF, RDFS, OWL, Topic Maps. Afin de bien argumenter l'utilité des ontologies, nous décrivons les composantes d'une ontologie. Après, nous présentons les différentes typologies, les outils de construction, les méthodologies et le cycle de vie d'une ontologie. Enfin, nous discutons le rôle joué par les ontologies dans les applications E-Learning.

4.2. Définition

4.2.1. Le Web Sémantique

Le Web Sémantique est une vision du futur Web dans lequel l'information est donnée dans un sens explicite facilitant ainsi aux machines le traitement et l'intégration des informations sur le Web. Le Web Sémantique sera construit sur la capacité de XML de définir des schémas de balisage personnalisés et sur la flexibilité de l'approche RDF pour représenter les données. Si les machines sont supposées faire des tâches de raisonnement utiles sur ces documents, le langage doit aller au-delà des sémantiques de base du RDF Schema. OWL a été conçu pour répondre à ce besoin pour un langage d'ontologie pour le Web. OWL est une partie d'une pile évolutive des recommandations W3C relativement au Web Sémantique (Boutemedjet, 2004). La Figure 4 ci-dessous (Benayache, 2005) présente les couches du Web Sémantique.

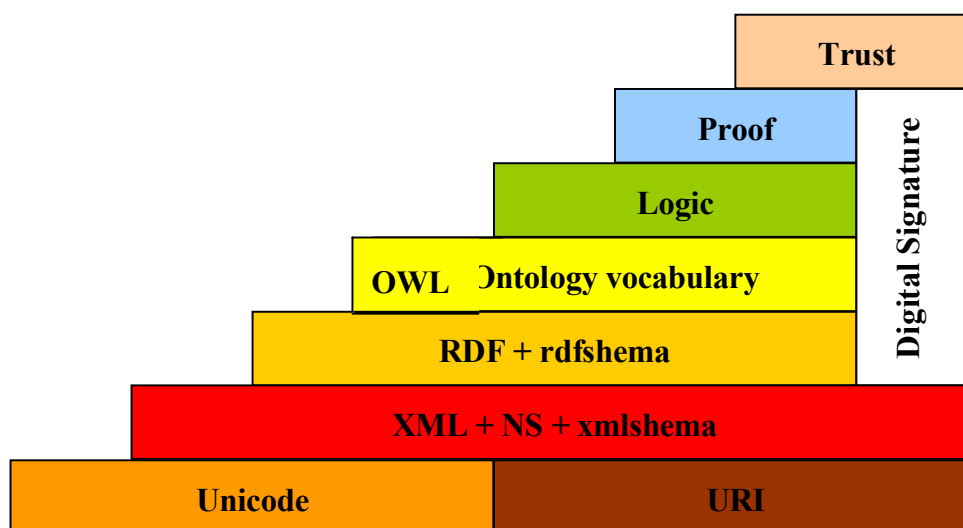


Figure 4 Les couches du Web Sémantique

XML (W3C, 1998) : fournit une surface syntaxique pour les documents structurés mais ne fournit aucune contrainte sémantique sur le sens de ces documents.

XML Schema (W3C, 2001) : est un langage pour restreindre la structure des documents XML et étendre aussi XML avec des types de données.

RDF (W3C, 2004, c) : est un modèle de données pour les objets 'ressources' et les relations entre eux, fournissant des sémantiques simples pour ce modèle de données qui peuvent être représentés en XML.

RDF Schema : est un vocabulaire pour décrire les propriétés et les classes des ressources RDF.

OWL : ajoute plus de vocabulaire pour décrire les propriétés et les classes entre autres, les relations entre les classes, cardinalité, égalité, typage de propriétés plus riche, caractéristiques des propriétés et les hiérarchies des propriétés et des classes.

4.2.2. L'ontologie

Ontologie : (O majuscule et singulier), (philosophie Grecque), *Ontos* : être *logos* : étude, l'étude de ce qui existe, c'est à dire l'ensemble des connaissances que nous l'avons sur le monde (Zaidi, 2008). Les ontologies : (o minuscule et pluriel) Introduites en (IA) vers les années 90, une ontologie est une spécification explicite d'une conceptualisation (Zaidi, 2008). Une ontologie est un réseau sémantique qui regroupe un ensemble de concepts décrivant complètement un domaine. Ces concepts sont liés les uns aux autres par des relations taxonomiques (hiérarchisation des concepts) d'une part, et sémantiques d'autre part (Khadir, 2007).

Les ontologies sont utilisées pour permettre aux experts d'un domaine de modéliser celui-ci afin qu'il soit compréhensible par les experts d'un autre domaine. Le modèle ainsi obtenu est utilisable informatiquement, c'est-à-dire qu'il est facile à représenter et à programmer. Une ontologie est disponible pour plusieurs sujets traitant du même domaine. En effet, il s'agit d'une analyse d'un domaine qui est adaptable à chaque cas particulier.

L'usage d'une ontologie appartient au milieu de l'intelligence artificielle. Il s'agit de générer des méthodes informatiques de compréhension de données réelles. L'ontologie définit des concepts appelés classes. Chaque classe contient des attributs et des restrictions de domaine de ces attributs. L'ensemble des instances de toutes les classes de l'ontologie constitue une base de connaissances (Khadir, 2007).

4.3. Objectif du web sémantique

L'expression de Web Sémantique, attribuée à Tim Berners-Lee (Berners-Lee, Hendler et Lassila, 2001) (fondateur et président du Consortium World Wide Web 'W3C'), fait d'abord référence à la vision du Web de demain comme un vaste espace d'échange de ressources entre humains et machines permettant une meilleure exploitation de masses de données disponibles sur le Web.

L'objectif est non pas de permettre aux machines de se comporter comme des êtres humains, mais de développer des langages pour représenter les informations d'une manière traitable par les machines, afin d'améliorer les rapports des utilisateurs avec le Web. En effet, la sémantique et la structure des données requièrent une représentation de la sémantique compréhensible et échangeable pas les machines. Comme l'écrit, en substance, Tim Berners-Lee, 'le Web Sémantique est ce qu'on obtiendra si on réalise le même processus de globalisation sur la représentation des connaissances que celui que le Web fait initialement sur l'hypertexte'.

Le Web Sémantique peut être défini comme un substrat supportant des fonctions avancées pour la collaboration (homme-homme, homme-machine, machine-machine), qui permet de partager des ressources et de raisonner sur le contenu de ces dernières (Berners-Lee, Hendler et Lassila, 2001). L'idée est de rendre explicite la sémantique des documents au travers de métadonnées ou d'annotations, afin de permettre aux agents logiciels d'effectuer des tâches de recherche et de sélection des ressources pour les utilisateurs.

Le Web Sémantique n'est pas un Web distinct mais bien un prolongement du Web que l'on connaît, dans lequel nous attribuons à l'information une signification clairement définie, ce qui permet aux ordinateurs et aux humains de travailler en plus étroite collaboration (Berners-Lee, Hendler et Lassila, 2001).

Les recherches actuellement réalisées dans le domaine du Web Sémantique s'appuient sur un existant riche venant de différents domaines. Par exemple, les systèmes de recherche en Intelligence Artificielle (IA), les systèmes de représentation et/ou l'ingénierie des connaissances ont permis d'étudier les problèmes liés à l'accès aux collections d'informations structurées, aux règles d'inférences et aux raisonnements automatiques bien avant le développement du Web. Cependant, l'application des résultats de ces recherches pose d'autres problèmes dus au changement du contexte de déploiement, le Web et ses dérivés (Internet, Extranet, Intranet), la nécessité d'un niveau élevé d'interopérabilité, les diversités des usages, les standardisations, ...etc.

Le défi du Web Sémantique est de fournir un langage (Benayache, 2005) :

- Qui exprime à la fois les données et les règles de raisonnement sur ces données,
- Qui permette aux règles de n'importe quel système de représentation des connaissances d'être transférées sur le Web.

Nous pouvons distinguer deux types d'approche pour le Web Sémantique (Benayache, 2005), l'une qualifiée de 'Web computationnellement sémantique' et l'autre de 'Web cognitivement sémantique'. La première concerne l'automatisation de la recherche au moyen d'agents logiciels. La deuxième s'intéresse à la structuration des contenus et vise une semi-automatisation de certaines tâches.

L'approche 'Web cognitivement sémantique' privilégie la problématique de l'indexation à celle de l'inférence. Elle permet néanmoins des inférences 'simples' à partir de représentations dont la 'sémantique opérationnelle' est plus faible que celle de représentations basées sur des langages formels supportant des traitements puissants (Benayache, 2005).

4.4. Les langages du Web Sémantique

Les travaux de standardisation sont bien avancés. Durant le 10 février 2004, OWL et RDF sont devenus des recommandations du W3C :

- RDF est utilisé pour représenter l'information et pour échanger la connaissance sur le Web.
- OWL est utilisé pour publier et partager les ensembles de termes (appelées les ontologies), en supportant les recherches Web avancées, la gestion de la connaissance et des agents logiciels (W3C 2004). (Boutemedjet, 2004) :

4.4.1. RDF

RDF est un langage d'assertions destiné à exprimer les propositions en utilisant des vocabulaires formels et précis, plus spécialement ceux exprimés en RDFS, pour accéder et utiliser le World Wide Web et il est destiné pour fournir les fondations de base pour les autres langages d'assertion avancés ayant les mêmes objectifs. Une assertion telle que définit sur le site du W3C est une expression prétendue être vraie et peut dépendre de plusieurs facteurs incluant les conventions sociales, les commentaires dans un langage naturel et des liens vers des documents fournissant du contenu (content bearing documents). La plupart de ce contenu inaccessible aux traitements des machines et est mentionné ici seulement pour envisager que les sémantiques formelles décrites dans le document de spécification n'ont pas l'intention de fournir une analyse complète du 'sens' (Boutemedjet, 2004).

Un document RDF est un ensemble de triplets de la forme < sujet, prédicat, objet >. Les éléments de ces triplets peuvent être des URI 'Universal Resource Identifier', des littéraux ou des variables. Cet ensemble de triplet peut être représenté de façon naturelle par un graphe plus précisément un multi-graphe orienté étiqueté où les éléments apparaissant comme sujet ou objet sont des sommets, et chaque triplet est représenté par un arc dont l'origine est son sujet et la destination est son objet. Ce document sera codé en machine par un document RDF/XML ou N3, mais est souvent représenté sous une forme graphique (Figure 4 en page 52).

4.4.2. OWL

Le langage OWL (Web Ontology Language) a été conçu pour être utilisé par les applications qui traitent le contenu de l'information au lieu de la présenter seulement aux êtres humains. OWL facilite grandement l'interopérabilité au niveau machine du contenu du Web plus que ce qui est déjà supporté par les XML, RDF et RDF Schema (RDF-S) en fournissant du vocabulaire supplémentaire avec des sémantiques formelles.

OWL peut être utilisé pour représenter le sens des termes dans des vocabulaires ainsi que les relations qui existent entre ces termes. Cette représentation des termes et de leurs interrelations est appelée ontologie. OWL est une révision du langage d'ontologie Web DAML+OIL (DAML : 'DARPA Agent Markup Language', OIL : 'Ontology Inference Layer') incorporant les leçons apprises de la conception et de l'application de DAML+OIL. OWL décrit la structure d'un domaine en terme de classes et de propriétés comme les approches orientées objets (Boutemedjet, 2004).

Bien que OWL est dérivé de DAML+OIL qui est équivalent à un langage très expressif de la logique des descriptions, en tant que tel n'est pas équivalent à aucune logique de description. Certaines caractéristiques font qu'il n'existe aucun algorithme d'inférence décidable avec toute cette puissance d'expression. Par exemple, les propriétés dans OWL peuvent être déclarées comme symétriques, une classe peut être traitée simultanément comme une collection d'individus ou comme un seul individu pour son propre compte. OWL Permet à une ontologie d'augmenter le sens du vocabulaire prédéfini (RDF et OWL). De ce fait OWL fournit des sous langages de plus en plus expressifs (OWL Lite, OWL DL et OWL Full) conçus pour faire un compromis entre son pouvoir expressif et son pouvoir de raisonnement (Boutemedjet, 2004).

OWL Lite : Supporte les utilisateurs qui ont besoin des hiérarchies de classifications et des caractéristiques de contraintes simples. Par exemple, lorsqu'il supporte les contraintes de cardinalité, il permet seulement les valeurs (0 et 1).

OWL DL : Supporte les utilisateurs qui demandent un maximum d'expressivité tout en maintenant la complétude (garantie de calculer toutes les conclusions) et la décidabilité (tous les calculs doivent finir en un temps fini). OWL DL contient tout les constructeurs du langage OWL mais sont utilisables avec des restrictions (par exemple, lorsqu'une classe peut être une sous classe de plusieurs autre classes, une classe ne peut être une instance d'une autre classe).

OWL Full : destiné aux utilisateurs qui demandent un maximum d'expressivité avec la liberté syntaxique de RDF sans aucune garantie de calculs. Par exemple, une classe peut être traitée comme une collection d'individus et en même temps peut être vue comme un seul individu. OWL Full permet aussi à une ontologie d'augmenter le sens du vocabulaire prédéfini (RDF et OWL).

Pour ces trois sous langages, seulement les deux premiers maintiennent les tâches d'inférence principales satisfiabilité/classement.

4.5. L'intérêt de l'ontologie

4.5.1. Un besoin générique

Pour faire tomber les barrières créées par des vocabulaires disparates. Un besoin existe de partager la signification de termes dans un domaine donné.

- Toute activité humaine spécialisée développe son propre jargon (langue de spécialité) sous la forme d'une terminologie et d'une conceptualisation associée spécifiques.
- L'existence de tels jargons entraîne des problèmes de compréhension et des difficultés à partager des connaissances entre les acteurs de l'entreprise, les services d'une entreprise, les entreprises d'une industrie, qui font des métiers différents.
- Fondamentalement, le rôle des ontologies est d'améliorer la communication entre humains, mais aussi entre humains et ordinateurs et finalement entre ordinateurs.

4.5.2. Une aide à la communication

Vers un vocabulaire standardisé de communication entre agents humains et aussi entre organisations.

L'existence de vocabulaires différents au sein d'une entreprise (ex : bureau d'études, bureau des méthodes) ou d'une industrie (ex : constructeur automobile, équipementier) constitue un frein à la collaboration et aux partenariats.

Les enjeux touchent donc directement la compétitivité de l'entreprise. Pour l'entreprise, l'ontologie sert à :

- améliorer la compréhension entre les employés,
- favoriser la diffusion des informations et leurs exploitations,
- promouvoir une nouvelle approche de conception des systèmes d'information (réutilisation de codes, interopérabilité des logiciels).

4.5.3. Une aide à la conception et à l'utilisation des systèmes d'information

Des apports pour l'ingénierie des systèmes d'information

- **Spécification** : Acquisition des connaissances : une ontologie peut aider à l'analyse des besoins et à définir les spécifications d'un système d'information.
- **Réutilisation** : Partage : une ontologie peut être, ou peut devenir suite à une traduction, un composant réutilisable et/ou partagé par plusieurs logiciels.
- **Fiabilité** : Maintenance : une ontologie peut servir à améliorer la documentation d'un logiciel et/ou à automatiser des vérifications de cohérence, réduisant les coûts de maintenance.
- **Interopérabilité** : en jouant le rôle d'un format d'échange, l'ontologie permet à des systèmes d'information, basés sur des paradigmes de modélisation et des langages d'implantation différents, de coopérer.

4.6. Composantes de l'ontologie

Une ontologie se compose de cinq types d'éléments (Khadir, 2007) : 1) les concepts ; 2) les relations ; 3) les fonctions ; 4) les axiomes ; et 5) les instances.

4.6.1. Les concepts

Aussi appelés termes ou classes de l'ontologie, correspondent aux abstractions pertinentes d'un segment de la réalité (le domaine du problème), retenues en fonction des objectifs qu'on se donne et de l'application envisagée pour l'ontologie.

4.6.2. Les relations

Traduisent les associations pertinentes existant entre les concepts présents dans le segment analysé de la réalité. Ces relations incluent les associations suivantes : Sous-classe-de (généralisation – spécialisation ou 'Is-a'), Partie-de (agrégation ou composition ou 'part-of'), Associée-à et Instance-de, ...etc.

4.6.3. Les fonctions

Sont des cas particuliers de relations dans lesquelles un élément de la relation, le nième (extrant), est défini en fonction des n-1 éléments précédents (intrants).

4.6.4. Les axiomes

Constituent des assertions acceptées comme vraies, à propos des abstractions du domaine traduites par l'ontologie.

4.6.5. Les modèles (ou instances)

Constituent la définition extensionnelle de l'ontologie.

4.7. Dimensions de classification

Les ontologies peuvent être classifiées selon plusieurs dimensions. Parmi celles-ci, nous en examinerons quatre (Psyché, Mendes et Bourdeau, 2003) : **1)** Objet de conceptualisation ; **2)** Niveau de détail ; **3)** Niveau de complétude et **4)** Niveau de formalisme de représentation.

4.7.1. Typologie selon l'objet de conceptualisation

Les ontologies classifiées selon leurs objet de conceptualisation par (Gómez-Pérez, 1999a), (Guarino, 1997b), (Mizoguchi, 1998), (Mizoguchi et Ikeda, 1996), (Van-Heijst, Schreiber et Wielinga, 1997), (Vanwelkenhuysen et Mizoguchi, 1994), (Vanwelkenhuysen et Mizoguchi, 1995), (Wielinga et Schreiber, 1993) selon (Psyché, Mendes et Bourdeau, 2003) , le sont de la façon suivante : **1)** Représentation des connaissances ; **2)** Supérieure/ Haut niveau ; **3)** Générique ; **4)** Domaine ; **5)** Tâche et **6)** Application.

1) Ontologie de représentation de connaissances

Selon (Gómez-Pérez, 1999) et (Van-Heijst, Schreiber et Wielinga, 1997), ce type d'ontologies regroupe les concepts (primitives de représentation) impliqués dans la formalisation des connaissances. Un exemple est l'ontologie de Frame qui intègre les primitives de représentation des langages à base de frames : classes, instances, facettes, propriétés/slots, relations, restrictions, valeurs permises, ...etc.

2) Ontologie supérieure ou de haut niveau

Selon (Guarino, 1997a) et (Sowa, 1995), cette ontologie est une ontologie générale. Son sujet est l'étude des catégories des choses qui existent dans le monde, soit les concepts de haute abstraction tels que: les entités, les événements, les états, les processus, les actions, le temps, l'espace, les relations, les propriétés. L'ontologie de haut niveau est fondée sur : la théorie de l'identité, la méréologie (theory of whole and parts role) et la théorie de la dépendance.

3) Ontologie générique

Selon (Gómez-Pérez, 1999a), (GómezPérez, 1999b) et (Van-Heijst, Schreiber et Wielinga, 1997), cette ontologie aussi appelée, méta-ontologies ou core ontologies, véhicule des connaissances génériques moins abstraites que celles véhiculées par l'ontologie de haut niveau, mais assez générales néanmoins pour être réutilisées à travers différents domaines. Elle peut adresser des connaissances factuelles (Generic domain ontology) ou encore des connaissances visant à résoudre des problèmes génériques (connaissances procédurales) appartenant ou réutilisables à travers différents domaines (Generic task ontology). Deux exemples de ce type d'ontologies sont : **1)** l'ontologie méréologique (Borst, 1997) contenant des relations, Partie-de et **2)** l'ontologie topologique contenant des relations, Associé-à.

4) Ontologie du domaine

Selon (Mizoguchi et Bourdeau, 2000), cette ontologie régit un ensemble de vocabulaires et de concepts qui décrit un domaine d'application ou monde cible. Elle permet de créer des modèles d'objets du monde cible. L'ontologie du domaine est une méta-description d'une représentation des connaissances, c'est-à-dire une sorte de méta-modèle de connaissance dont les concepts et propriétés sont de type déclaratif. La plupart des ontologies existantes sont des ontologies du domaine.

Selon Mizoguchi, l'ontologie du domaine caractérise la connaissance du domaine où la tâche est réalisée. Dans le contexte de la formation à distance, un domaine serait par exemple : le téléapprentissage.

5) Ontologie de tâches

Selon (Mizoguchi et Bourdeau, 2000), ce type d'ontologies est utilisé pour conceptualiser des tâches spécifiques dans les systèmes, telles que les tâches de diagnostic, de planification, de conception, de configuration, de tutorat, soit tout ce qui concerne la résolution de problèmes.

Elle régit un ensemble de vocabulaires et de concepts qui décrit une structure de résolution des problèmes inhérente aux tâches et indépendante du domaine. Selon (Mizoguchi et Bourdeau, 2000), l'ontologie de tâche caractérise l'architecture computationnelle d'un système à base de connaissances qui réalise une tâche. Deux exemples d'utilisation de l'ontologie de tâche dans le domaine de l'éducation sont les suivants : **1)** l'ontologie de formation par ordinateur - Computer Based Training Ontology (Jin et al, 1999) : qui régit un ensemble de concepts spécifiques à un système d'apprentissage inhérent à des ontologies de tâche ; et **2)** l'ontologie des objectifs d'apprentissage - Learning Goal Ontology (Inaba et al, 2000) : qui décrit les rôles des apprenants et des agents dans le cadre d'un apprentissage collaboratif.

6) Ontologie d'application

(Maedche, 2002) cette ontologie est la plus spécifique. Une métadonnée est littéralement une donnée sur une donnée. Plus précisément, c'est un ensemble structuré d'informations décrivant une ressource quelconque. Elles décrivent des concepts dépendants d'un domaine et d'une tâche particulière. Ces concepts correspondent souvent au rôle joué par les entités d'un domaine dans une tâche précise. Une ontologie est constituée d'un ensemble de concepts qui ont des relations de différents types entre eux.

4.7.2. Typologie selon le niveau de détail de l'ontologie

Par rapport au niveau de détail utilisé lors de la conceptualisation de l'ontologie en fonction de l'objectif opérationnel envisagé pour l'ontologie, deux catégories au moins peuvent être identifiées (Psyché, Mendes et Bourdeau, 2003) :

1) **Granularité fine** : correspondant à des ontologies très détaillées, possédant ainsi un vocabulaire plus riche capable d'assurer une description détaillée des concepts pertinents d'un domaine ou d'une tâche. Ce niveau de granularité peut s'avérer utile lorsqu'il s'agit d'établir un consensus entre les agents qu'ils utiliseront ;

2) **Granularité large** : correspondant à un vocabulaire moins détaillé comme par exemple dans les scénarios d'utilisation spécifiques où les utilisateurs sont déjà préalablement d'accord à propos d'une conceptualisation sous-jacente. Les ontologies de haut niveau possèdent une granularité large, compte tenu que les concepts qu'elles traduisent sont normalement raffinés subséquemment dans d'autres ontologies de domaine ou d'application.

4.7.3. Typologie selon le niveau de complétude

Bachimont propose la classification sur trois niveaux comme suit (Psyché, Mendes et Bourdeau, 2003) :

Niveau 1 - Sémantique : Tous les concepts (caractérisés par un terme/libellé) doivent respecter les quatre principes différentiels : 1) Communauté avec l'ancêtre; 2) Différence (spécification) par rapport à l'ancêtre; 3) Communauté avec les concepts frères (situés au même niveau); 4) Différence par rapport aux concepts frères (sinon il n'aurait pas lieu de le définir). Ces principes correspondent à l'engagement sémantique qui assure que chaque concept aura un sens univoque et non contextuel associé. Deux concepts sémantiques sont identiques si l'interprétation du terme/libellé à travers les quatre principes différentiels aboutit à un sens équivalent.

Niveau 2 - Référentiel : Outre les caractéristiques énoncées au niveau précédent, les concepts référentiels (ou formels) se caractérisent par un terme/libellé dont la sémantique est définie par une extension d'objets. L'engagement ontologique spécifie les objets du domaine qui peuvent être associés au concept, conformément à sa signification formelle. Deux concepts formels seront identiques s'ils possèdent la même extension (ex : les concepts d'étoile du matin et d'étoile du soir associés à Vénus).

Niveau 3 - Opérationnel : Outre les caractéristiques énoncées au niveau précédent, les concepts du niveau opérationnel ou computationnel sont caractérisés par les opérations qu'il est possible de leur appliquer pour générer des inférences (engagement computationnel). Deux concepts opérationnels sont identiques s'ils possèdent le même potentiel d'inférence.

4.8. Les outils de construction d'ontologie

Il existe deux types de construction 'les outils dépendants de formalisme de représentation et les outils indépendants de formalisme de représentation' (Khadir, 2007) :

4.8.1. Les outils dépendants de formalisme de représentation

1) Ontolingua

Ontolingua est un serveur d'édition d'ontologies : une ontologie est directement exprimée dans un formalisme également nommé Ontolingua, qui constitue en fait une extension du langage KIF 'Knowledge Interchange Format'. Le langage Ontolingua utilise des classes, des relations, des fonctions, des objets (instances) et des axiomes pour décrire une ontologie.

Une relation (ou une classe) peut contenir des propriétés (contraintes) nécessaires et suffisantes qui définissent la relation (ou la classe). En plus de ces fonctionnalités offertes par le langage Ontolingua, le serveur Ontolingua offre la possibilité d'intégrer les ontologies Ontolingua, ce qui permet la construction modulaire des ontologies. Il y a trois différentes possibilités d'intégrer les ontologies Ontolingua :

- **Inclusion** : Une ontologie inclut et utilise les définitions d'autres ontologies.
- **Restriction** : l'ontologie importe les définitions depuis d'autres ontologies et les rend plus spécifiques.
- **Raffinement polymorphe** : nous redéfinissons une définition importée depuis n'importe quelle ontologie.

2) OntoSaurus

OntoSaurus de l'Information Science Institute de l'Université de Southern California. L'OntoSaurus est composé de deux modules : un serveur utilisant LOOM comme langage de représentation des connaissances, et en un serveur de navigation créant dynamiquement des pages HTML qui affichent la hiérarchie de l'ontologie; le serveur utilise des formulaires HTML pour permettre à l'utilisateur d'éditer l'ontologie. Nous pouvons représenter les concepts, la taxonomie des concepts, les relations entre les concepts, les fonctions, les axiomes et les instances.

3) WebOnto

WebOnto du Knowledge Media Institute de l'Open University, est une application Web pour naviguer et développer collaborativement les ontologies OCML 'Operational Conceptual Modelling Language'. WebOnto supporte la navigation collaborative, la création et l'édition d'ontologies sur le Web. Les ontologies WebOnto sont implémentées dans le langage OCML. Le langage OCML est une combinaison des frames et de la logique de premier ordre et permet de représenter les concepts, la taxonomie des concepts, les relations, les fonctions, les axiomes et les instances. WebOnto distingue quatre types d'ontologies : ontologie de domaine, ontologie de tâche, ontologie de méthode, et ontologie d'application. WebOnto supporte l'inclusion d'ontologie au moyen des interfaces graphiques. En ce qui concerne l'édition collaborative d'ontologie, WebOnto est le seul outil qui procure cette fonctionnalité permet aux ingénieurs des connaissances de tenir des discussions sur les changements et les mises à jour des ontologies lors d'édition ou de navigation, en mode synchrone et asynchrone.

4) OilEd

OilEd (Oil Editor) est un éditeur d'ontologies utilisant le formalisme OIL. Il est essentiellement dédié à la construction de petites ontologies dont nous pouvons ensuite tester la cohérence à l'aide de FACT, un moteur d'inférences bâti sur OIL. Le modèle de connaissance a été adapté depuis OIL à DAML+OIL et maintenant il va être adapté à OWL.

4.8.2. Les outils indépendants de formalisme de représentation

1) Protégé2000

Protégé2000 est une interface modulaire permettant l'édition, la visualisation, le contrôle (vérification des contraintes) d'ontologies, l'extraction d'ontologies à partir de sources textuelles, et la fusion semi-automatique d'ontologies. Le modèle de connaissances sous-jacent à Protégé2000 est issu du modèle des frames et contient des classes (concepts), des slots (propriétés) et des facettes (valeurs des propriétés et contraintes), ainsi que des instances des classes et des propriétés. Protégé2000 se veut compatible avec le modèle OKBC 'Open Knowledge Base Connectivity' mais en diffère sur certains points. Par exemple, un frame, qui est soit une classe, soit un slot, soit une facette, est instance d'une classe et une seule dans Protégé2000 ce qui n'est pas toujours le cas dans OKBC, où le même frame peut d'ailleurs représenter plusieurs objets conceptuels.

Protégé2000 autorise la définition de méta-classes, dont les instances sont des classes, ce qui permet de créer son propre modèle de connaissances avant de bâtir une ontologie.

2) ODE et WebOde

L'outil ODE 'Ontology Design Environment' permet de construire des ontologies au niveau connaissance, comme le préconise la méthodologie METHONTOLOGY. L'utilisateur construit son ontologie dans un modèle de type frame, en spécifiant les concepts du domaine, les termes associés, les attributs et leurs valeurs, les relations de subsomption. L'ontologie opérationnelle est alors générée en utilisant les formalismes ONTOLINGUA ou FLOGIC. WEBODE est l'adaptation d'ODE pour le Web.

3) OntoEdit

OntoEdit (Ontology Editor) est également un environnement de construction d'ontologies indépendant de tout formalisme.

Il permet l'édition des hiérarchies de concepts et de relations et l'expression d'axiomes algébriques portant sur les relations, et de propriétés telles que la généralité d'un concept. Des outils graphiques dédiés à la visualisation d'ontologies sont inclus dans l'environnement. OntoEdit intègre un serveur destiné à l'édition d'une ontologie par plusieurs utilisateurs. Un contrôle de la cohérence de l'ontologie est assuré à travers la gestion des ordres d'édition. Enfin, un plug-in nommé ONTOKICK offre la possibilité de générer les spécifications de l'ontologie par l'intermédiaire de questions de compétences.

4.9.Méthodologie et cycle de vie d'une ontologie

Une Méthodologie est un ensemble de procédures de travail et les étapes qui décrivent le pourquoi et le comment de la conceptualisation. Bien qu'il n'existe pas de méthodologie ayant fait l'unanimité, plusieurs ont été proposées, les plus importantes sont :

1. Méthodologie de (Uschold et Grüninger, 1996).
2. Methontology : proposée par (Fernandez-Lopez, Gómez-Perez et Juristo, 1997).
3. Méthodologie de (Guarino et Welty, 2000a et 2000b).
4. Méthode Archonte (ARCHitecture for ONTological Elaborating) proposée par (Bachimont, 2000).

Nous détaillons une seule méthodologie 'Uschold et King', c'est celle que nous l'avons utilisée dans notre système.

4.9.1. (Uschold et King, 1995)

Ils ont proposé à USC 'University of Southern California' (Khadir, 2007) une méthode de construction d'ontologie qui est basée sur l'expérience acquise lors du développement de l'ontologie d'Entreprise 'The enterprise ontology'.

Pour construire une ontologie conformément à l'approche de Uschold et King, les étapes suivantes doivent être respectées : identification des objectifs et du contexte de l'ontologie - construire l'ontologie - capture de l'ontologie - codage de l'ontologie - intégration d'ontologies existantes - évaluer l'ontologie et documenter l'ontologie.

Première étape : identification des objectifs et du contexte : le but de cette étape est de clarifier le pourquoi de la construction de l'ontologie et les utilisations prévues, la finalité (réutilisée, partagée, utilisée comme une partie d'un KB, ...etc.), les utilisateurs potentiels de l'ontologie.

Deuxième étape : construction de l'ontologie : cette étape est divisée en trois activités : **1)** capture de l'ontologie, **2)** codage de l'ontologie, **3)** intégration d'ontologies existantes.

- Première activité : capture de l'ontologie : Capture de l'ontologie indépendamment d'un langage de représentation (conceptualisation) : identification des concepts et des relations clés, produire en langue naturelle les définitions précises et non ambiguës pour les concepts et relations, identification des termes dénotant les concepts et les relations, pour ainsi essayer d'arriver à un agrément. Pour la catégorisation et la capture de l'ontologie, (Uschold et Grüninger, 1996) proposent trois approches :

- **Approche descendante** : partir de concepts abstraits que l'on spécialise en concepts plus spécifiques.
- **Approche ascendante** : partir de tous les concepts spécifiques que l'on généralise en concepts abstraits.
- **Approche intermédiaire** : les concepts se structurent autour de concepts intermédiaires, ni trop généraux, ni trop spécifiques.

- Deuxième activité : codage de l'ontologie : Cette activité implique deux tâches : **a)** représentation explicite de la conceptualisation (ex. classe, entité, relation), **b)** écrire le code dans un langage formel : Prolog, KL-One, OIL, CG, OWL...

- **Troisième activité : intégration d'ontologies existantes** : Cette activité fait référence à comment utiliser les ontologies qui existent. Cette activité peut être faite en parallèle avec les activités de capture et/ou de codage. Par exemple ; The Frame Ontology peut être réutilisé pour modéliser les ontologies de domaine qui utilise l'approche basée sur les frames.

Troisième étape : évaluation de l'ontologie : confrontation de l'ontologie : objectifs, logiciels, utilisateurs.

Quatrième étape : documentation de l'ontologie : essentielle pour l'acceptation de l'ontologie.

Malgré le nombre important de méthodes et de démarches proposées, à l'heure actuelle nous en comptons une trentaine, aucune n'a pu s'imposer. Ces méthodologies peuvent porter sur l'ensemble du processus et guider l'ontologiste sur toutes les étapes de la construction d'ontologies.

L'essentiel dans toutes ses étapes la, c'est la possibilité de mettre l'ontologie en évolution, donc mettre à jour l'ontologie via un cycle de vie (Voir Figure 5 ci-dessous, (Gandon, 2006)).

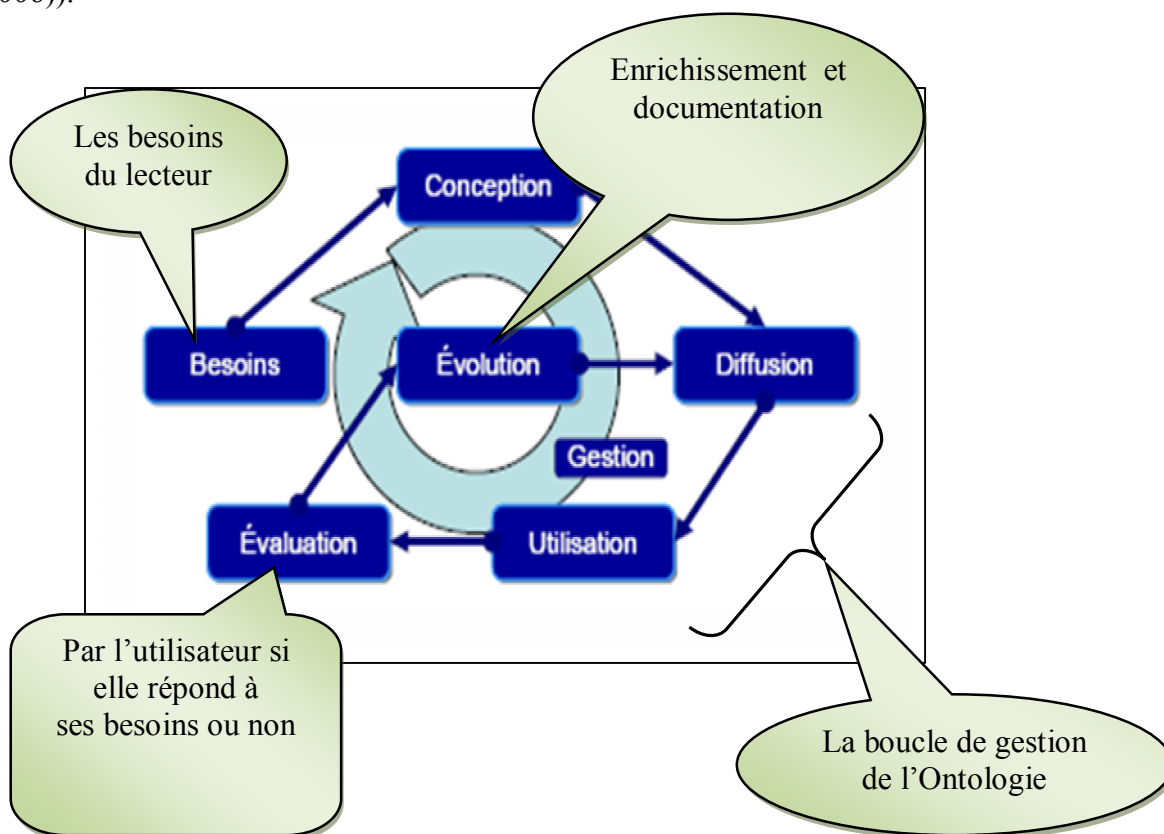


Figure 5 Le cycle de vie de l'Ontologie

Les activités liées à une ontologie peuvent être regroupées en trois catégories (Benayache, 2005) :

- des activités de gestion de projet : planification, contrôle, assurance qualité ;
- des activités de développement : spécification, conceptualisation, formalisation ;
- des activités de support : évaluation, documentation, gestion de la configuration.

4.10. Rôle des ontologies pour les applications E-Learning

(Uschold et Grüninger, 1996) considère que les ontologies varient suivant trois dimensions :

Le degré de formalisation de la représentation : qui varie continument depuis le rigoureusement formel jusqu'au hautement informel ;

L'objectif opérationnel : communication entre utilisateurs, interopérabilité entre systèmes, application à un problème d'ingénierie comme la réutilisation de composants et résolution de problèmes ;

Le sujet : domaine de connaissance, connaissances de raisonnement, connaissances liées au modèle de représentation.

Le degré de formalisation doit être choisi en fonction des contextes d'usage de l'ontologie, car il peut y avoir contradiction entre, d'une part, un besoin de formalisme lié au partage des connaissances et à la réutilisabilité, et, d'autre part, la nécessité de produire une ontologie compréhensible par les utilisateurs.

4.11. Utilité des ontologies pour le E-Learning

Une ontologie est toujours construite en fonction de son usage. Dans le cas d'un usage E-Learning, l'utilisation d'ontologies peut intervenir pour :

- **La communication :** elle permet de créer un 'vocabulaire' standardisé et facilite la communication entre les différents acteurs de la formation, entre les systèmes et entre les acteurs et les systèmes.
- **L'interopérabilité (communication entre systèmes) :** elle permet de soutenir la conception des systèmes de formation E-Learning. Le degré de formalité requis doit être rigoureusement formel.

- **Le partage et la réutilisation des informations pédagogiques :** l'utilisation des ontologies permet la mise en correspondance des contenus sémantiques des objets pédagogiques utilisés dans différents systèmes. Elle facilite la réutilisation des objets pédagogiques par d'autres systèmes, au moyen d'une traduction automatique.
- **L'indexation et la recherche d'information pédagogique :** les ontologies peuvent être utilisées comme méta-descripteur pour décrire le contenu sémantique associé aux objets.
- **Mise à jour des informations pédagogiques :** les ontologies peuvent contribuer à rendre le processus de mise à jour des objets pédagogiques qu'elles décrivent, plus facile et plus efficace.
- **Les connaissances du domaine E-Learning :** l'utilisation des ontologies permet d'améliorer le processus d'acquisition de connaissances lors de la construction d'une formation E-Learning. Cela se traduit par une meilleure organisation des objets pédagogiques et des connaissances du domaine.
- **La spécification de notions :** les ontologies peuvent être utilisées comme moyen de spécification des notions à appréhender et des relations entre notions d'une formation donnée.

4.12. Les ontologies pour les applications E-Learning

Nous décrivons dans cette section quelques exemples d'utilisation d'ontologies dans des systèmes de formation E-Learning (Psyché, Mendes et Bourdeau, 2003) :

Le projet QBLS : pour l'étudiant, QBLS 'Question Based Learning System' est un outil d'aide à la résolution de questions de Travaux Dirigés (TD) à partir d'éléments de cours annotés sémantiquement. Pour l'enseignant, QBLS est vu non seulement comme un outil permettant de diffuser son support électronique sous une forme facilitant la navigation, mais également comme un moyen d'inciter les étudiants à aller chercher activement les connaissances.

Dans cette optique, QBLS utilise une ontologie du domaine limitée à un thésaurus. Les concepts du domaine sont des instances des classes modélisant les types de savoirs clés du cours.

Le projet DIVA-BCTA : les objectifs de ce projet sont : 1) de créer une base de connaissances sur le téléapprentissage par le développement d'une ontologie du domaine, et 2) de regrouper l'expertise sur le téléapprentissage dans des objets pédagogiques référencés en fonction de cette ontologie. L'ontologie du domaine sert de système d'indexation et de classification des objets pédagogiques. Par ailleurs, une ontologie de tâche permet de définir différents cas ou scénarios d'utilisation de ces objets, identifiés grâce à leurs référencements en fonction de l'ontologie du domaine. Le rôle de l'ontologie de tâche est de fournir des réponses aux requêtes des utilisateurs de la base.

Le projet de repérage et de visualisation du modèle de l'apprenant 'VIUM' : Dans le cadre de ce projet, un outil de repérage et de visualisation, VIUM, a été conçu pour permettre à l'utilisateur de sélectionner un concept central sur l'écran. Une ontologie est utilisée pour s'assurer que les concepts les plus proches sémantiquement sont visibles. Cette sélection de concepts rendus visibles est une partie essentielle de la visualisation qui assiste les apprenants dans l'exploration de domaines comprenant des centaines de concepts. La tâche particulière ici, est de montrer à un utilisateur ce qu'une ontologie computationnelle permet d'inférer à partir d'informations. C'est le cas pour l'auto-évaluation par l'apprenant.

Le projet de composition dynamique de documents pédagogiques: ce projet utilise une ontologie pour repérer des briques d'informations pertinentes. Le contenu des briques est validé grâce à l'ontologie du domaine. Ainsi, la vérification de la pertinence consiste à associer à chaque brique un ou plusieurs fragments significatifs de cette ontologie, traduisant la sémantique contenue dans la brique. Le repérage d'information est réalisé en fonction d'objectifs spécifiés par un apprenant. Parallèlement, l'ontologie est utilisée pour organiser les briques dont le contenu est pertinent par rapport aux attentes d'un apprenant. Les auteurs du projet ont testé deux méthodes différentes : l'une en fonction de grammaires formelles, et l'autre par association de deux ontologies (du domaine et pédagogique). Celle ayant démontré le plus de souplesse est la méthode basée sur les ontologies.

Le projet Web Sémantique et E-Learning : l'objectif de ce projet est de proposer des concepts, des méthodes et des outils sur les environnements informatiques pour l'apprentissage humain mettant en œuvre l'expertise de ses membres dans les domaines de l'ingénierie des données et des connaissances, de l'extraction de connaissances et de la recherche de documents et d'informations. Il relève par là même à la fois de la recherche théorique et de la R&D. Parmi les travaux réalisés dans le cadre de ce projet, nous trouvons :

- **Aide à l'enseignant pour la composition de cours :** L'objectif est d'aider l'enseignant dans sa phase de préparation de cours. Il s'agit de modéliser et synchroniser les ressources pédagogiques disponibles sur le web, en vue de concevoir, d'évaluer et de réviser un cours. Ce travail permet la création assistée de cours par l'enrichissement d'une ontologie de concepts pédagogiques.
- **Parcours personnalisés pour le E-Learning :** L'objectif est de pouvoir enrichir les connaissances sur la navigation dans des parcours de formation. Pour cela, les auteurs modélisent le site web en utilisant une ontologie du domaine (représentée en XML) et relient les pages du site à cette ontologie. Ensuite, ils catégorisent les apprenants pour pouvoir obtenir des connaissances ciblées. Ces informations vont permettre par catégories d'apprenants : d'optimiser les parcours, de ne présenter que les informations pertinentes et d'aider à la révision du site.

Le projet de développement de système conseiller : ce projet utilise une ontologie de tâches pour identifier et définir les différents composants de la tâche de conseil et les différentes notions qui vont être prises en considération dans le futur système. Deux types de tâche ont été identifiés : 1) celle de l'apprenant qui suit une formation via un environnement d'apprentissage diffusé sur le Web ; 2) celle d'un concepteur qui construit un cours à l'aide d'un système conseiller accessible via le Web.

Le projet IMAT: l'objectif du projet IMAT 'Integrating Manuals And Training' est de proposer des modèles et des outils pour permettre à un concepteur de documents pédagogiques d'intégrer des parties de documents techniques dans les documents de formation qu'il produit. Il est focalisé sur la maintenance d'équipements complexes et utilise une ontologie pour indexer des briques de documents. L'idée est la composition dynamique de documents indépendamment des contextes d'utilisation : Technique ; Médiatique ; Domaine ; ou Pédagogique. L'ontologie pédagogique utilisée permet d'attribuer des rôles aux briques ainsi que des attributs qualitatifs et des contraintes d'organisation des briques dans le document final.

Le projet de développement de l'environnement de conception de curriculum CREAM et de tutorat CREAM-TOOL : ce projet utilise quatre ontologies constituant la base du curriculum : une ontologie de capacités, une ontologie d'objectifs, une ontologie de ressources, et une ontologie de liens.

Ces ontologies servent à décrire un module de capacités d'apprentissage, un module d'objectifs d'apprentissage, un module de ressources didactiques et des liens de couplage entre les trois modules précédents. Elles sont construites à partir d'alphabets (Klausmeier, etc.), de vocabulaires d'évaluation (Merrill, Winograd, etc.), de taxonomies (Cagné, Bloom, etc.) et d'un ensemble de concepts (capacités, objectifs, ressources, liens, etc.) reliés entre eux (relation d'analogie, de généralisation, d'agrégation, de déviation, de pré-requis, de prétexte, de constitution, de similarité, etc.). Dans le cadre de ce projet, un environnement de type tuteur pour la génération automatique de curriculum CREAM et de cours CREAM-TOOLS a été conçu.

Le projet de développement d'environnements pour favoriser l'apprentissage collaboratif : Ce projet porte sur la construction d'une ontologie pour représenter les concepts communs des théories d'apprentissage collaboratif.

Plusieurs théories démontrent les avantages de l'apprentissage collaboratif (i.e., l'apprentissage par l'observation, l'apprentissage autoréguler, la cognition distribuée, la théorie de flexibilité cognitive, la théorie socioculturelle, etc.).

En sélectionnant une théorie parmi celles-ci et en formant un groupe d'apprenants, nous pouvons nous attendre à un apprentissage collaboratif efficace s'appuyant sur cette théorie. Selon Inaba (2001) (Psyché, Mendes et Bourdeau, 2003), les environnements d'apprentissage collaboratif doivent être conçus en fonction de ces théories, c'est-à-dire qu'ils doivent contenir une ontologie du domaine de l'apprentissage collaboratif. Un système contenant une telle ontologie fournit des fonctionnalités de médiation de l'apprentissage aux apprenants, leur permettant de cette façon de construire leurs connaissances.

4.13. Conclusion

Le Web Sémantique apparaît comme une technologie prometteuse pour implémenter le E-Learning. L'objectif est d'améliorer les rapports des utilisateurs avec le Web. Le Web Sémantique peut être défini comme un substrat supportant des fonctions avancées pour la collaboration (homme-homme, homme-machine, machine-machine), qui permet de partager des ressources et de raisonner sur le contenu de ces dernières (Berners-Lee, Hendler et Lassila, 2001). Une des caractéristiques principales est la compréhension partagée basée sur un squelette d'ontologie (Benayache, 2005).

Le Web Sémantique peut représenter une plate-forme adéquate pour implémenter des systèmes E-Learning. Ainsi, nous avons parlé de l'exploitation des ontologies dans le domaine de formation E-Learning. Les ontologies apparaissent désormais comme une clé pour la manipulation automatique de l'information au niveau sémantique.

L'adoption du concept d'ontologie par l'ingénierie des connaissances a donné lieu à l'émergence d'une multiplicité de méthodes et d'outils de modélisation et de mises en œuvre. Malgré les critiques que nous pouvons apporter à l'ensemble de ces méthodes et outils, il est nécessaire d'utiliser une méthodologie ou une démarche à suivre pour le développement d'une ontologie, afin de permettre un meilleur cadrage des connaissances. Selon (Farquhar, Fikes et Rice, 1996) dans (Benayache, 2005), quelque soit la méthodologie utilisée, le processus de construction d'une ontologie est une collaboration qui réunit des experts du domaine de connaissance, des ingénieurs de la connaissance, voire les futurs utilisateurs de l'ontologie.

Les résultats d'exploitation des ontologies au sein des applications et environnement E-Learning, affichent un potentiel riche et multiple. Cela laisse à penser que leur place au sein des systèmes E-Learning ne peut que croître. Si l'ensemble des projets utilisant des ontologies ne vise, pour le moment, que la gestion des informations et connaissances pédagogiques, l'application des ontologies à terme pourrait permettre le développement de systèmes E-Learning capables non seulement de gérer les connaissances mais aussi de raisonner sur ces connaissances.

L'idée de tous ce que nous avons vu c'est de pouvoir utiliser des techniques issues de l'ingénierie des connaissances pour construire une ontologie de domaine et une autre pédagogique, avec un but bien précis. Le but de l'ontologie pédagogique est bien entendu l'adaptation du contenu de la formation extrait à partir du Web aux profils des apprenants. Dans le chapitre suivant, nous parlons sur les différentes techniques d'adaptation dans les systèmes E-Learning.

5. Les techniques d'adaptation dans le E-Learning

5.1. Introduction

Le présent chapitre décrit les notions de base de l'adaptation. Ensuite, il présente l'apport entre les systèmes éducatifs en ligne intégrant des dispositifs d'adaptation. Ces systèmes sont capables d'adapter selon les critères suivants : Les données de l'utilisateur, les données d'usage, et l'environnement. Puis, il donne une vision globale sur les paramètres d'adaptation qui sont relatifs à l'utilisateur et ceux qui en sont indépendants, les différentes méthodes et techniques d'adaptation et ses éléments.

Le point qui rend E-Learning plus pratique que les méthodes traditionnelles d'apprentissage est que le E-Learning peut être adressé au nombre maximal de participants qui peuvent avoir une diversité dans les styles d'apprentissage, les préférences et les besoins. Pour cela, un système E-Learning adaptatif doit avoir la qualité de la formation [(Akoumianakis, 2011) et (Lee, Hsieh et Hsu, 2011)].

En plus, E-Learning offre des avantages tels que la réduction des coûts et du temps, ce qui prouve que la réalisation et la certification sont des éléments essentiels des initiatives de la formation tels que : l'interactivité et la capacité d'offrir des contenus cohérents qui sont nécessaires en fonction des demandes (Kruse, 2004). Donc, ces environnements d'apprentissage en ligne veulent arriver à une certaine satisfaction aux exigences des apprenants, qui ont un rôle central dans l'apprentissage. Pour cela, ces environnements sont devenus de plus en plus populaires [(Chen, 2009), (Lee, Hsieh, et Hsu, 2011), (Monova-Zheleva, 2005), (Rosmalen et al, 2006) et (Wang et al, 2010)].

Le point principal d'un système E-Learning populaire est qu'il doit prendre en considération l'étude de l'apprenant pour déterminer ses préférences, ses intérêts et ses comportements de navigation afin d'offrir une formation adaptée [(Chen, 2009) et (Wang et al, 2011)]. Cette considération met en évidence une idée de la gestion d'un système d'apprentissage adaptatif (Brusilovsky, 1999). Cette méthode est considérée comme une alternative aux méthodes traditionnelles et ouvre la porte à de nouvelles approches dans le développement des systèmes éducatifs adaptatifs [(Akoumianakis, 2011), (Caravantes et Galán, 2011), (Peterson, Rayner et Armstrong, 2009), (Lee, Hsieh, et Hsu, 2011), (Surjono, 2009), (Wang et al, 2010) et (Wang et al, 2011)].

Beaucoup d'efforts ont été accomplis dans le domaine des systèmes éducatifs adaptatifs pour offrir un bon modèle apprenant dans les systèmes d'apprentissage (Nguyen et Do, 2008). La plupart de ces travaux sont sur les méthodes d'apprentissage des apprenants afin d'acquérir plus de détails sur leurs besoins [(Liegle et Janicki, 2006), (Magoulas, Papanikolaou, et Grigoriadou, 2003), (Stach, Cristea et De Bra, 2004) et (Yang et Wu, 2009)]. Il ya deux approches générales différentes dans l'adaptation des contenus d'apprentissage [(Olfman, Mandviwalla, 1994) et (Papanikolaou et *al*, 2002)]. La première approche vise à adapter le contenu d'apprentissage ayant des besoins spéciaux. La seconde se concentre sur la fourniture du contenu d'apprentissage le plus approprié aux besoins de l'apprenant. La première s'appelle l'adaptation au niveau du contenu et la deuxième s'appelle l'adaptation au niveau des liens. Aucune de ces approches ne peut être préférable de l'autre.

Sur la base de ces approches, plusieurs projets de recherches ont donné une nouvelle méthodologie pour avoir un contenu approprié. Certains de ces projets sont orientés vers l'extension du contenu standard d'apprentissage pour améliorer la qualité du processus d'apprentissage. Un groupe affirme que les normes actuelles ne supportent pas un système adaptatif, de sorte qu'ils doivent être modifiés selon certains respects [(Lu et Hsieh, 2008), (Rey-Lopez et *al*, 2009), (Sampson, Karagiannidis et Cardinali, 2002)].

En réponse au fait que les normes de métadonnées de contenus d'apprentissage sont en quelque sorte insuffisante pour certaines applications, un groupe de chercheurs ont essayé de les remplacer par l'ontologie qui est considérée comme l'un des piliers du Web sémantique [(Chi, 2009), (Jovanovic et *al*, 2007), (Lee, Tsai et Wang, 2008), (Shih, Yang et Tseng, 2009), (Verbert et *al*, 2005), (Wang et Hsu, 2006), (Yang, 2006) et (Zitko et *al*, 2009)]. Les chercheurs ont développé des systèmes basés sur le Web sémantique en créant une ontologie utilisée pour les thèmes du cours. Cette méthode de modélisation permet une interaction entre les apprenants et les systèmes comme ayant été effectuées dans les recherches par [(Chi, 2009), (Jovanovic et *al*, 2007), (Lee, Tsai et Wang, 2008), (Shih, Yang, et Tseng, 2009), (Verbert et *al*, 2005), (Wang & Hsu, 2006), (Yang, 2006) et (Zitko et *al*, 2009)].

Il y a quelques études qui ont utilisé un système multi-agents dans l'apprentissage adaptatif [(Canales et *al*, 2007) et (Chen, 2008)]. Quatre types d'agents existent dans ce système proposé : le premier est un agent de gestion du contexte, le second fait le choix du contenu, le troisième concerne l'organisation du contenu et le quatrième fait la présentation du contenu.

5.2. Définitions

5.2.1. Adaptation

Action d'adapter ou de s'adapter à quelque chose : adaptation aux circonstances (LAROUSSE). Le terme adaptation ou adapter est composé de : ad et apt. En latin, ad est un préfixe qui exprime le rapprochement, la proximité, l'addition, le renforcement, le commencement. Pour apte, le dictionnaire propose les synonymes, capable, compétent, propice, convenable, avoir les qualités requises et, pour adapte, ce même dictionnaire propose, ajuste, met en accord, arrange... bref s'adapter, aller vers plus d'aptitude ! (ADAPTE, 2013).

En informatique, L'adaptation à l'usager (terme officiel désignant la customisation en anglais), est l'adaptation d'un matériel (ou d'un logiciel) fabriqué en série aux exigences particulières d'un usager (ENMI12, 2012).

5.2.2. Adaptabilité

C'est la capacité de s'adapter à de nouveaux milieux ou à de nouvelles situations (LAROUSSE).

L'adaptabilité d'un système concerne sa capacité à réagir selon le contexte, et selon les besoins et préférences des utilisateurs (Bastien, Leulier et Scapin, 1998).

Justification : Plus les façons d'effectuer une même tâche sont diverses, plus les chances que l'utilisateur puisse choisir et maîtriser l'une d'entre elles, au cours de ses apprentissages, sont importantes. Il faut donc fournir à l'utilisateur des procédures, options, et commandes différentes lui permettant d'atteindre un même objectif. Par ailleurs, une interface ne peut convenir à la fois à tous ses utilisateurs potentiels. Pour qu'elle n'ait pas d'effets négatifs sur l'utilisateur, cette interface doit, selon les contextes, s'adapter à l'utilisateur (Scapin et Bastien, 1997).

5.2.3. Adaptation, Adaptabilité et Adaptativité

Il existe dans la littérature une différence entre adaptabilité et adaptativité : 'L'adaptabilité est la capacité du système de s'adapter aux personnalisations expressément demandées par l'utilisateur, tandis que l'adaptativité désigne sa capacité à répondre aux besoins de l'utilisateur sans une intervention explicite de sa part' (Moisuc, 2006).

En termes d'action, l'adaptabilité renvoie à un processus dit statique, basé essentiellement sur des connaissances disponibles ou acquises par le système avant que les interactions utilisateur-système ne soient engagées. Ainsi, les adaptations sont réalisées lors de l'initialisation du système qui se présente dans une version adaptée à l'utilisateur. De plus, les connaissances utilisées par le système semblent inchangées au cours d'utilisation (edutechwiki, 2006).

Quant à l'adaptativité, elle a une vision plus dynamique du processus d'adaptation. En effet, les connaissances sont acquises ou modifiées par le système au cours des interactions. En suite, le système procède à des adaptations pendant que l'utilisateur interagit avec lui (Stephanidis et *al*, 1998). Par ailleurs, l'utilisateur prend le contrôle de ses actions (initiation, proposition, sélection, exécution) sur le système dans l'adaptabilité alors qu'il n'exerce aucun contrôle sur le système dans l'adaptativité. (Villanova-Oliver, 2002, page 20-21).

L'adaptation, qu'elle soit adaptabilité ou adaptativité, nécessite la définition d'un ou plusieurs modèles d'utilisateur. Le concepteur doit avoir une certaine représentation de l'utilisateur (edutechwiki, 2006). Entre l'usage des concepteurs et celle des utilisateurs, les résultats montrent souvent des écarts considérables. C'est une des principales raisons qui pousse à considérer d'autant plus fortement l'utilité du concept d'adaptation (edutechwiki, 2006).

(Bastien, Leulier et Scapin, 1998) évoquent que : 'l'adaptabilité d'un système concerne sa capacité à réagir selon le contexte, et selon les besoins et préférences des utilisateurs'. Mais ils vont plus loin en ajoutant que cette adaptabilité se réfère à la 'flexibilité', et à la 'prise en compte de l'expérience de l'utilisateur'. Cela signifie que l'utilisateur, au gré de sa pratique et de son niveau, serait sensé obtenir des fonctions et des commandes différentes : dans le but d'éviter des effets négatifs sur l'utilisateur, en fonction de son contexte et de ses besoins (edutechwiki, 2006). En d'autres termes, c'est au système de s'adapter à l'utilisateur, selon le niveau d'expertise, en lui offrant des moyens disponibles non seulement pour personnaliser les fonctions, mais aussi pour répondre aux exigences souhaitées par une multitude d'actions.

Dans le domaine de la formation, plus un environnement d'apprentissage offre des caractéristiques d'adaptation, plus l'apprentissage peut être, individualisé. Autrement dit, un logiciel adaptable est en mesure de répondre à un besoin spécifique de la part de l'usager en cours d'utilisation.

Pour illustrer notre argumentation, nous prendrons l'exemple d'une des possibilités existantes sur un logiciel de messagerie instantanée qui est MSN. Selon (Dico du Net, 2013), ce logiciel est : 'Plus rapide que le courrier électronique, c'est un outil idéal pour les conversations et la solution parfaite lorsque vous ne pouvez pas être présent en personne'. Nous ne nous attarderons pas sur le logiciel en lui-même, mais seulement sur le concept d'adaptabilité. Aussi, nous l'aurons compris, l'idée de la messagerie instantanée est bien de pouvoir correspondre rapidement avec d'autres usagers. Mais si, en cours de communication l'utilisateur A désire parler à l'utilisateur B, en incluant d'autres utilisateurs tels que C et D dans la même conversation, le besoin alors ne sera plus le même. Et c'est justement à partir de ce point de vue que nous pouvons reconnaître que la fonction 'inviter quelqu'un à rejoindre la conversation' que A va activer (en cliquant sur l'icône y relatif) devient une source d'adaptabilité pour nos usagers. Par contre, sous l'angle de l'adaptativité, il s'agit davantage d'une possibilité que le logiciel effectue de lui-même, selon les habitudes ou les fonctions employées par l'utilisateur. Cela revient à dire que c'est le système qui agit sans que personne ne lui ait demandé quoique ce soit de concret.

Mais pour clarifier davantage cette idée, nous citerons les termes de (Dillenbourg, 1994), lorsqu'il dit : 'Un système capable de décider quel concept ou quelle compétence enseigner sur la base du comportement de l'apprenant est potentiellement plus adaptatif qu'un système dans lequel le curriculum est fixé une fois pour toute'. Par exemple, il se peut qu'une configuration ait été effectuée une fois pour que le système s'en souvienne et qu'alors ce dernier propose à l'usager la ou les fonctions que ce dernier avait tantôt souhaité.

Les logiciels ayant les caractéristiques d'adaptabilité et d'adaptativité peuvent être qualifiés d'intelligents'. Nous parlons alors d'intelligence artificielle'. Toutefois, ces deux concepts apportent des éléments différents : au contraire de l'adaptabilité, l'adaptativité d'un logiciel permet à l'utilisateur d'entrer plus facilement dans l'environnement qui est lui est offert car il n'a pas besoin de faire des manipulations pour agir sur la situation, c'est l'outil qui anticipe et réagit.

L'impact en terme d'apprentissage ne sera alors pas le même. Dans le cas d'un jeu éducatif, par exemple un micro monde, l'utilisateur apprendra des choses sans s'en rendre compte. Mais rien ne dit qu'il pourra transférer ces compétences acquises 'à son insu' dans d'autres situations car il n'aura pas pris conscience des procédures mises en œuvre dans son apprentissage.

5.2.4. Adaptation et Pédagogie

Nous venons de constater que les utilisateurs n'ont pas tous le même niveau, ni les mêmes besoins les uns des autres. De fait, pour les utilisateurs inexpérimentés, il est recommandé de fournir un guidage 'étape par étape', contrairement aux utilisateurs 'experts'. Par exemple, certains logiciels préparent d'avance des séquences pour les 'novices' et des fonctions supplémentaires pour les experts. De sorte que cette mise à disposition centrée sur l'utilisateur profite à long terme aux deux usagers (edutechwiki, 2006).

Selon les tenants du cognitivisme et du constructivisme, l'usage d'un artefact est un bon moyen pour viser un apprentissage. Et dans un environnement informatisé d'apprentissage, les interactions sont alors soutenues par une ou des composantes informatiques. Mais il ne s'agit pas d'introduire un élément informatisé pour que se dessine une stratégie pédagogique. Nous voulons dire par là que pour qu'il y ait la construction d'un apprentissage, il faudrait que le système proposé à un groupe d'apprenants puisse réellement susciter l'intérêt, la motivation et permettre de se forger de nouvelles représentations.

Dès lors, la nécessité d'identifier des outils accessibles et adéquats, efficaces et efficaces, renvoie directement à l'utilité de proposer une activité informatisée adaptée aux apprenants et au contexte d'apprentissage.

5.2.5. Flexibilité

Réfère aux moyens disponibles à l'utilisateur pour personnaliser son interface de façon à prendre en compte ses stratégies de travail et/ou ses habitudes et les exigences de ses tâches (Bastien et Scapin, 1993).

5.2.6. Prise en compte de l'expérience

Ce sont les différents moyens disponibles pour prendre en compte le niveau d'expérience de l'utilisateur (expérimenté, débutant, occasionnel) (Bastien et Scapin, 1993).

5.3. Les systèmes d'enseignement en ligne et l'adaptation

En accord avec l'approche proposée par Edmonds, nous distinguons trois types de systèmes intégrant des dispositifs d'adaptation 'Adapté, Adaptable et Adaptatif' (Edmonds, 1981) et (Benadi, 2004) :

5.3.1. Un système adapté

C'est un système dans lequel l'adaptation est l'œuvre du concepteur lui-même. Le système prend en compte, une fois pour toutes, un profil d'utilisateur ou un groupe d'utilisateurs définis préalablement à sa mise en place. Les techniques d'adaptation sont appliquées durant la phase de conception du système (par 'conception participative' par exemple) et l'adaptation ne peut, donc, pas être propre à chaque individu.

5.3.2. Un système adaptable

Est un système qui peut être modifié, à tout moment, sur demande explicite de l'utilisateur qui le consulte. Celui-ci peut choisir l'aspect physique des éléments de l'interface à l'aide par exemple de tableaux de bord. Il peut également demander la création de macro-commandes (comme sous Windows). Dans ce type de système, ce sont les utilisateurs qui saisissent leurs préférences, qui les enregistrent dans un modèle qui, par la suite, n'est modifié que sur une nouvelle demande explicite de l'utilisateur.

5.3.3. Un système adaptatif

Est un système qui s'adapte de lui-même, à chaque instant, aux besoins et les habitudes de l'utilisateur. Des mécanismes de suivi (tracking) des comportements de l'utilisateur permettent de connaître les besoins de ce dernier en fonction de son environnement, de son état psychologique et de ses connaissances. La mise à jour du modèle utilisateur est réalisée par le système lui-même, par observation des actions et des réactions de l'utilisateur (interaction avec le système).

Dans la première catégorie de systèmes (système adapté), l'adaptation est l'œuvre du concepteur. Les utilisateurs (ou une classe d'utilisateurs) n'interviennent pas mais sont identifiés a priori à travers leurs besoins. Le système peut uniquement adapter son comportement durant l'exécution à partir d'un échantillon de modèles de comportements prédéfinis (notions de stéréotypes) qui peuvent plus ou moins pertinents (Benadi, 2004).

Dans les deux autres types de systèmes, les besoins et les habitudes de l'utilisateur sont prises en compte, mais la différence entre l'adaptabilité et l'adaptativité réside dans le fait que cette dernière ne demande pas à l'utilisateur de se préoccuper des adaptations et de la mise à jour de ses préférences, qui doivent, donc, être déduites des actions qu'il entreprend.

Dans de tels systèmes, l'utilisateur n'a pas à apprendre comment adapter son interface ou à mémoriser d'éventuels langages de spécifications pour opérer les modifications nécessaires à l'adaptation. Il peut, donc, concentrer son énergie sur le fonctionnement du logiciel et l'accomplissement de sa tâche en espérant que le système comprenne correctement ses objectifs et ses désirs et mémorise ces habitudes (Benadi, 2004). Cette adaptativité, évoluant d'une façon autonome par rapport à l'utilisateur, présente des avantages manifestés mais aussi des inconvénients comme le risque de désorientation. s'explique par le non respect de l'un des principes de base de l'ergonomie logicielle qui précisant que dans un système bien conçu les mêmes actions doivent toujours provoquer les mêmes réactions.

5.4. Critères d'adaptation

Selon (Brusilovsky, 2001) et (Fenineche, 2007) les systèmes d'enseignement en ligne sont capables d'adapter selon les critères suivants : Les caractéristiques de l'utilisateur, les données d'usage, et l'environnement.

5.4.1. Adaptation aux caractéristiques de l'utilisateur

Les caractéristiques de l'utilisateur pris en compte sont : ses références ou acquis (background), ses connaissances du sujet, ses objectifs (goals), ses préférences et son expérience dans l'hypermédia. Brusilovsky a cité deux autres caractéristiques dans (Brusilovsky, 2001) qui sont : Les choix de l'utilisateur (intérêts) et les traits individuels.

5.4.1.1. Connaissances

Les connaissances du domaine sont variables pour un utilisateur donné, ce qui signifie qu'un système hypermédia doit prendre en considération les changements de l'état des connaissances et met à jour le modèle de l'utilisateur avec les nouvelles données. Ces connaissances sont représentées par des modèles dont les plus populaires sont : Le modèle 'OVERLAY' et le modèle 'PUGGY' (Chabert-Ranwez, 2000).

5.4.1.1.1. Le modèle par recouvrement (Overlay)

Ce modèle est généralement constitué d'un ensemble de concepts du domaine organisé en réseau sémantique, qui définit la structure du domaine. Il consiste à modéliser l'apprenant en rapportant ses connaissances supposées par recouvrement de la modélisation du domaine et est construit au fur et à mesure que l'apprenant avance dans le cours (nœuds visités).

Chaque représentation de concepts est alors stockée avec une estimation de la connaissance de l'apprenant pour ce concept : binaire (su, pas su), qualitative (bon, moyen, pauvre) ou quantitative (la probabilité que l'utilisateur connaisse ce concept). Le modèle 'Overlay' paraît adapté pour représenter toute la connaissance de l'utilisateur nécessaire pour séquencer un curriculum personnalisé en hypertexte et fournir une aide à la navigation.

5.4.1.1.2. Le modèle Buggy

Le système Buggy a été mis en place pour déterminer la combinaison de procédures correctes et incorrectes la plus proche possible du comportement de l'élève et de tenter d'y remédier. Cette méthode a été largement reprise par la suite dans des formes plus évoluées de diagnostic d'erreurs. En fonction du diagnostic obtenu, une nouvelle leçon est générée pour corriger les erreurs ou les confusions de l'apprenant (Delectre, 2000).

5.4.1.2.Objectifs (goals ou tasks)

Est une caractéristique reliée au contexte du travail de l'utilisateur dans l'hypermédia, elle peut être une réponse à la question : pourquoi l'utilisateur utilise-t-il le système et que veut-il atteindre ? Cette caractéristique change d'une session à une autre et plusieurs fois lors d'une session.

Dans certains systèmes, il est utile de distinguer les objectifs locaux ou à niveau bas qui changent souvent et les objectifs généraux ou de haut niveau qui sont stables. Par exemple, pour les systèmes éducatifs, l'objectif d'apprentissage est de haut niveau, par contre la résolution d'un problème est un objectif de bas niveau qui peut changer plusieurs fois lors d'une session.

5.4.1.3.Références (background) et expérience

Deux autres caractéristiques de l'utilisateur qui sont similaires aux connaissances de l'utilisateur mais différentes sur le plan fonctionnel, ce sont les références et l'expérience dans l'hypermédia. Les références de l'utilisateur sous entendent toute l'information liée à l'expérience passée par ce dernier en dehors du sujet des systèmes d'enseignement en ligne. Cela inclue la profession de l'utilisateur, l'expérience de travail dans les domaines relatifs au sujet. L'expérience quant à elle signale à quel point l'utilisateur est familier avec la structure de l'hypermédia et avec quelle facilité il navigue dans le système ?

5.4.1.4.Préférences

Pour différentes raisons, un utilisateur peut préférer certains nœuds et liens sur d'autres et quelques parties d'une page par rapport à d'autres. Ces préférences peuvent être absolues ou relatives c'est à dire dépendantes du nœud courant, de l'objectif, ou du contexte courant en général. Différemment des autres composantes, les préférences ne peuvent être déduites par le système, l'utilisateur informe le système directement ou indirectement (avec un simple feed back) de ses préférences. Elle est plus liée à l'adaptabilité qu'à l'adaptativité. La différence c'est que les systèmes d'enseignement en ligne peuvent généraliser les préférences de l'utilisateur et les appliquer pour l'adaptation dans de nouveaux contextes. Les préférences sont souvent représentées et calculées numériquement avec des moyens spéciaux.

Cette représentation a des avantages : elle ouvre la possibilité pour combiner plusieurs modèles utilisateurs. Les modèles de groupe accumulent les préférences d'un groupe spécifique d'utilisateurs. Un modèle de groupe est un bon modèle de départ pour un nouveau membre du groupe. Il est aussi important pour le travail collaboratif.

5.4.1.5.Choix

Les Choix (intérêts) de l'utilisateur ont été pris en charge avec le développement des systèmes hypermédias en recherche d'information sur le web. Ces derniers ont tenté de modéliser les choix à long terme et utiliser ceux-ci parallèlement avec le but de recherche à court terme pour améliorer le filtrage des informations.

5.4.1.6.Traits individuels

Les traits individuels englobent les facteurs de personnalité (introverti, extraverti), les facteurs cognitifs, les styles d'apprentissage. Comme les références de l'utilisateur, les traits individuels sont des caractéristiques stables qui ne peuvent être changées ou qui changent après une longue durée. Ces traits ne sont pas tirés d'un simple entretien mais après des tests psychologiques.

Plusieurs théorie sont été développées relatives aux traits psychologiques d'un apprenant qui agissent sur sa manière d'apprendre, mais il n'existe pas (à notre avis) un consensus sur la détermination de traits psychologiques qui peuvent être pris en considération dans le domaine de l'apprentissage et plus spécifiquement l'apprentissage à distance.

Dans (Hong et Kinshuk, 2004), nous trouvons, quatre styles principaux, selon la théorie de Felder-Silverman 'Le modèle Felder-Silverman', élaboré par Richard Felder, professeur de génie chimique, et Linda K. Silverman, psychologue scolaire, a été formulé à l'origine dans un contexte d'ingénierie et scientifique et s'inspire des modèles de (Kolb, 1984) :

- Active (do it) /reflective (Think about it)
- Sensing (learn facts)/intuitive (learning concepts)
- Visual (Require pictures) / Verbal (require reading or lecture)
- Sequential (step by step) / global (big picture)

Cette théorie est dotée d'un questionnaire qui peut être implémenté sur une plate forme de E-Learning et les résultats peuvent aisément être exploités par un environnement adaptatif.

Dans (Santally, 2005), nous retrouvons une classification des traits psychologiques selon quatre classes qui sont : le style cognitif, le contrôle cognitif, le style d'apprentissage et les performances :

- **Style cognitif**

Organisation de l'information (serial/holist) : Les 'holists' ont une vue globale et des interprétations larges de leurs environnements, tandis que les 'serialists' sont analytiques, ils se concentrent sur les détails avant de voir globalement leur environnement.

Assemblage des informations (Visuel/Auditif/Kinesthésique) : Cet instrument a été développé par Barbe Swassing et Milone (Barbe, Swassing et Milone, 1988 et 1979) qui décrit l'apprenant comme visuel, Auditif ou Kinesthésique. Les personnes avec une préférence visuelle tendent à montrer une plus grande capacité d'analyse et intégrer l'information visuelle. Un apprenant Auditif, préférerait traiter l'information sous forme écrite ou parlée. Les étudiants Kinesthésiques préfèrent traiter l'information par des moyens tactiles tels que des médias interactifs.

- **Contrôles Cognitifs**

Dépendance / indépendance du domaine : Les individus dépendants du domaine perçoivent des objets en général, tandis que les étudiants indépendants du domaine se concentrent sur des parties de l'objet.

Flexibilité cognitive / resserrement 'Constriction' cognitif : La flexibilité cognitive détermine la capacité d'un apprenant d'ignorer les distractions de son environnement tandis qu'il se concentre sur de l'information actuelle. Un individu qui a une grande flexibilité ne serait pas aussi facilement distrait que quelqu'un qui est 'constrictif'.

- **Style d'apprentissage (Théoricien/Activiste, Réflecteur/Pragmatiste)**

Les gens apprennent de manières différentes selon ce qu'on appelle 'style d'apprentissage'. La recherche et la pratique en matière d'éducation ont démontré que l'apprentissage peut être meilleur quand le processus d'instruction s'adapte aux divers styles d'apprentissage des apprenants. Le questionnaire de (Mumford et Honey, 1992) définit quatre styles d'apprentissage : théoricien, activiste, réflecteur et pragmatiste. Les personnes avec des préférences d'activiste, sont bien équipées pour l'expérience. Les personnes avec une approche 'réflecteur' sont bien équipées pour la réflexion. Les personnes avec des préférences de théoricien, sont bien disposées pour tirer des conclusions. Et enfin, les personnes avec des préférences de pragmatiste, avec leur tendance d'aimer les choses pratiques, sont très aptes pour la planification.

- **Performances**

La capacité d'assimilation et d'apprentissage d'un apprenant d'un objet pédagogique constitue un facteur important lors de la décision si un apprenant a satisfait l'étape requise en termes d'exécution d'une activité pour passer à un niveau différent.

5.4.2. Adaptation à l'usage

Comprend les données qui concernent l'interaction de l'utilisateur avec un système et qui ne peuvent être résolues avec les caractéristiques de l'utilisateur.

5.4.3. Adaptation à l'environnement

L'adaptation à l'environnement de l'utilisateur est un nouveau type d'adaptation. Les systèmes d'enseignement en ligne courants suggèrent quelques techniques pour s'adapter à la localisation de l'utilisateur, et à la plate-forme utilisée. Une adaptation simple à la plate-forme (hardware, software, bande passante) souvent implique la sélection du type du matériel et média pour présenter le contenu. Quant à la localisation, un nombre d'applications explore des techniques d'adaptation qui tiennent en compte la localisation de l'utilisateur et la direction de ses mouvements.

5.5. Méthodes et techniques d'adaptation des systèmes éducatifs

5.5.1. Paramètres d'adaptation

Nous entendons par paramètres d'adaptation les éléments sur lesquels repose la prise de décision du processus d'adaptation. Ces éléments vont être utilisés soit comme des déclencheurs de l'adaptation ou des sources de données. Nous distinguons entre deux types de paramètres qui peuvent être présents : ceux qui sont relatifs à l'utilisateur et ceux qui en sont indépendants (Hocine et *al*, 2011) :

5.5.1.1. Modèle utilisateur:

Il s'agit de l'ensemble de variables et métriques décrivant les caractéristiques de l'utilisateur dans le système. Ces caractéristiques peuvent être des données représentant les préférences de l'utilisateur, son état attentionnel, ses émotions et/ou ses compétences déduites à partir de l'analyse de ses traces. Ces données sont stockées dans le profil de l'utilisateur qui sera utilisé comme paramètre de processus d'adaptation.

5.5.1.2. Paramètre non-utilisateur ou variable système

Ce paramètre représente les variables propres au système et qui ne dépendent pas du modèle utilisateur. A titre d'exemple, nous pouvons citer les paramètres liés à la configuration matérielle et logicielle du système hôte (par exemple, l'accélération graphique des rendus 3D)

5.5.2. Modèle d'adaptation

L'adaptation peut être implémentée dans le système sous forme d'un module qui interagit avec le système pour modifier son comportement et sa structure. Ce module utilise un modèle qui réifie la politique d'adaptation souhaitée qui peut être soit implicite ou explicite :

Implicite : Dans ce cas les procédures d'adaptation se retrouvent éparpillées et étroitement liées aux différents composants du système. Il serait dans ce cas difficile de séparer dans les instructions (ou le code source) les éléments qui incombent à l'adaptation des autres aspects.

Explicite : Dans ce cas la technique d'adaptation utilise des modèles explicites comme un moteur de règles logiques, une matrice de décision, des algorithmes d'IA tels que les réseaux de neurones ou bien des algorithmes d'apprentissage automatique.

Il sera possible d'identifier et de séparer les aspects d'adaptation des autres aspects dans l'implémentation du système. L'objectif de ce critère est d'identifier la nature du modèle utilisé par la technique d'adaptation. Dans le cas d'un modèle explicite, ce critère identifie également le type de modèle utilisé (machine à états, règles logiques, modèle d'apprentissage automatique, ...etc.)

5.5.3. Eléments d'adaptation

Les hypermédias selon (Chikh, 2004), consistent en un ensemble de nœuds connectés par des liens où chaque nœud contient quelques informations locales et un nombre de liens aux autres pages. Ils peuvent aussi inclure un index et une carte globale qui fournissent des liens à toutes les pages accessibles. Les éléments d'adaptation sont le contenu des pages et les liens à partir de ces pages, pages d'index et cartes.

Il existe donc deux niveaux d'adaptation : le premier, appelé présentation adaptative, correspond au niveau du contenu et le second, appelé support de navigation adaptative, correspond au niveau des liens. Dans (Villanova-Oliver, 2002) et (Fenineche, 2007), une troisième dimension est ajoutée : adaptation de la présentation. Cette expression ne fait cependant pas référence à l'adaptation de la présentation telle que considérée ci-dessus mais il s'agit d'adapter les caractéristiques de mise en page (i.e. la forme ou l'apparence).

5.5.4. Méthodes et techniques

Une méthode d'adaptation selon (Villanova-Oliver, 2002) et (Fenineche, 2007), désigne une idée conceptuelle, une manière d'envisager une adaptation. Une technique d'adaptation correspond aux moyens mis en œuvre pour implémenter une méthode.

Une méthode peut être implémentée grâce à différentes techniques, et, inversement, une technique peut être utilisée pour différentes méthodes. Les différentes techniques et méthodes de la présentation adaptative et le support de navigation adaptative sont présentées dans (Chikh, 2004), (Brusilovsky, 1996) et (Brusilovsky, 2001).

5.5.4.1. Présentation adaptative

Il s'agit d'adapter le contenu d'une page visitée par un utilisateur particulier à son profil, un certain nombre de méthodes et techniques pour la présentation adaptative ont été élaborées ces dernières années.

5.5.4.1.1. Méthodes de la présentation adaptative

- **Les explications supplémentaires**

L'objectif des méthodes d'adaptation du contenu est de dissimuler à l'utilisateur quelques parties d'information au sujet d'un concept particulier qui ne sont pas pertinentes par rapport à ses connaissances. Par exemple, les détails à bas niveau d'un concept donné peuvent être dissimulés à l'utilisateur ayant un niveau de connaissances très faible parce qu'il ne peut pas comprendre ces détails. Au contraire, les explications supplémentaires, demandées généralement par les novices pour comprendre un concept donné, peuvent être dissimulées à l'utilisateur ayant un bon niveau de connaissance au sujet de ce concept parce qu'il n'aura jamais besoin de celles-ci.

- **Les explications pré-requises et les explications comparatives**

Elles changent l'information présentée au sujet d'un concept donné, en fonction du niveau de connaissance de l'utilisateur des concepts qui lui sont liés. La première méthode, explications pré-requises, est basée sur les liens 'pré-requis' entre les concepts. L'idée est qu'avant l'explication d'un concept, le système explique d'abord tous ses concepts 'pré-requis' qui ne sont pas suffisamment connus par l'utilisateur et qui sont nécessaires à sa compréhension. La seconde méthode, explications comparatives, est basée sur les liens de similarité entre les concepts. Si un concept, similaire au concept présenté, est connu par l'utilisateur, celui-ci aura une explication comparative qui fait apparaître les similarités et les différences entre le concept actuel et le concept lié.

- **Les variantes d'explications**

Avec cette technique, le système enregistre différentes variantes pour quelques parties du contenu de la page. Ainsi, l'utilisateur obtient la variante qui correspond à son modèle utilisateur.

- **Le tri d'informations au sujet d'un concept donné**

Une méthode aussi intéressante que les précédentes, qui peut tenir compte du niveau de connaissances et du profil de l'utilisateur, est le tri de fragments d'information au sujet d'un concept donné. L'information la plus proche du profil et des connaissances de l'utilisateur est placée devant.

5.5.4.1.2. Techniques de la présentation adaptative

❖ Le texte conditionnel

Avec cette technique, toutes les informations possibles au sujet d'un concept sont divisées en plusieurs portions d'informations (chunks). A Chaque portion est associée une condition au niveau des connaissances de l'utilisateur représentées dans le modèle utilisateur. Quand nous présentons l'information au sujet d'un concept, le système présente uniquement les portions où la condition est vraie.

❖ L'expansion de texte 'stretchtext'

La technique du texte allongé 'stretchtext' est un type spécial d'hypertexte. Dans les textes allongés, un texte lié peut remplacer le mot activé (ou la phrase) étendant ainsi le texte de la page courante. Une caractéristique importante des techniques du texte allongé 'stretchtext' est qu'ils laissent l'utilisateur adapter le contenu d'une page particulière, en fonction de ses connaissances et préférences.

❖ Les variantes de fragments et les variantes de page

La méthode des variantes d'explication peut être implémentée par les techniques des variantes de fragments et des variantes de pages. Avec la technique des 'variantes de pages', un système garde deux ou plusieurs variantes de la même page avec différentes présentations du même contenu. Lors de la présentation d'une page, le système sélectionne la variante de page selon le stéréotype de l'utilisateur courant. 'Les variantes de fragments' est une implémentation qui permet aux utilisateurs ayant différentes connaissances d'un concept particulier d'avoir différentes explications de ce concept.

❖ Les frames

Avec cette technique toutes les informations au sujet d'un concept particulier sont représentées sous forme de 'frame'. Les slots d'un frame peuvent contenir différentes variantes d'explication d'un concept. Des règles spéciales de présentation sont utilisées pour décider quels slots devraient être présentés à un utilisateur particulier et dans quel ordre. Dans leurs parties conditionnelles, ces règles peuvent faire référence non seulement au niveau des connaissances de l'utilisateur du concept présenté, mais aussi à n'importe quelle caractéristique représentée dans le modèle utilisateur.

5.5.4.2.Support de navigation adaptative

L'idée du support de navigation adaptative est d'aider les utilisateurs à trouver leurs chemins dans l'hypermédia en adaptant la présentation des liens aux objectifs, connaissances et autres caractéristiques de l'utilisateur. Avant de présenter les différentes méthodes et techniques, il faut savoir distinguer les quatre types de présentation des liens (Brusilovsky, 1996) : les liens locaux non contextuels ; les liens contextuels ou liens 'hypertextes réels' qui dépendent du contenu de la page ; les liens à partir d'index et tables de matières et les liens dans les cartes locales et dans les cartes de l'hyperespace global.

5.5.4.2.1. Méthodes du support de navigation adaptative

- **Le guidage global**

Il peut être fourni dans les systèmes hypermédias quand les utilisateurs ont quelques objectifs informationnels globaux (information qui est contenue dans un ou plusieurs nœuds de l'hyperespace) par la visualisation de l'information voulue. L'objectif des méthodes de guidage global est d'aider l'utilisateur à trouver le plus court chemin vers l'objectif informationnel avec une pénibilité minimale.

- **Le guidage local**

L'objectif des méthodes de guidage local est d'aider l'utilisateur à faire une étape de navigation par la suggestion des liens les plus pertinents à suivre à partir du nœud courant. Les méthodes de guidage local ne prévoient pas un objectif global, pour fournir le guidage. Ils font des suggestions selon les préférences, connaissances et profil de l'utilisateur quelle que soit l'importance du domaine. Par exemple, des méthodes pertinentes du guidage local, utilisées dans les hypermédias éducatifs sont : le tri des liens et le conseil direct selon les connaissances de l'utilisateur.

- **Le support d'orientation locale**

L'objectif des méthodes de support d'orientation locale est d'aider l'utilisateur à comprendre qu'est ce qui est autour et quelle est sa position actuelle. Les systèmes hypermédias adaptatifs implémentent les supports d'orientation locale de deux façons différentes : Fournir des informations supplémentaires au sujet des nœuds disponibles à partir du nœud courant (utiliser la technologie d'annotation) et limiter le nombre d'opportunités de navigation pour diminuer la surcharge cognitive et laisser les utilisateurs se concentrer sur l'analyse des liens les plus pertinents (utiliser la technologie de masquage des liens).

▪ **Le support d'orientation globale**

Le but des méthodes de support d'orientation globale est d'aider l'utilisateur à comprendre la structure de tout l'hyperespace et sa position absolue à l'intérieur de celui-ci. Cet objectif est généralement réalisé en fournissant des points de repère visibles et des cartes globales qui peuvent directement aider l'utilisateur dans l'orientation globale et en fournissant des visites guidées pour aider l'utilisateur graduellement à comprendre l'hyperespace. La fonction des annotations comme point de repère : tant qu'un nœud garde la même annotation quand l'utilisateur le voit de différentes positions dans l'hyperespace, l'utilisateur peut plus facilement reconnaître les nœuds qu'il a rencontrés avant et comprend la position courante. Masquer les liens réduit le volume de l'hyperespace visible et peut simplifier l'orientation et l'apprentissage.

▪ **La gestion des vues personnalisées**

La gestion des vues personnalisées est un nouvel objectif des systèmes hypermédia adaptatifs. Les vues personnalisées constituent une façon pour organiser une place de travail électronique pour les utilisateurs qui ont besoin d'un accès à une partie raisonnablement petite de l'hyperespace pour leur travail quotidien.

5.5.4.2.2. Techniques du support de navigation adaptative

❖ **Le conseil direct**

Cette technique permet d'indiquer à l'utilisateur le meilleur nœud prochain à visiter selon ses objectifs et les autres paramètres représentés dans son modèle. Pour fournir un guidage direct, le système peut esquisser visuellement le lien vers le meilleur nœud ou créer un lien dynamique supplémentaire (généralement appelé 'suivant') vers ce nœud.

❖ **Le tri adaptatif**

L'idée est de trier tous les liens d'une page particulière selon le modèle utilisateur : le lien le plus près du sommet est le plus important. Le tri adaptatif a une applicabilité limitée : il peut être utilisé avec des liens non contextuels, mais il peut à peine être utilisé pour les index et tables de matière et ne peut jamais être utilisé avec les liens contextuels et les cartes. Un autre problème avec le tri adaptatif est que cette technique rend l'ordre des liens non stable. En même temps quelques recherches récentes ont montré que la stabilité de l'ordre des options dans les menus est importante pour les novices.

❖ Le masquage des liens

L'idée du support de navigation par masquage est de restreindre l'espace de navigation en masquant les liens aux pages jugées 'non pertinentes' (ne correspond pas aux objectifs actuels de l'utilisateur ou présente des connaissances que l'utilisateur n'est pas préparé à comprendre). Cette technique protège les utilisateurs de la complexité de l'hypermédia non restreint et réduit la surcharge cognitive. Le masquage des liens a une applicabilité très étendue : elle peut être utilisée avec tous les types de liens non contextuels, d'index et des cartes par des boutons de visibilité réelle ou éléments de menus, et avec les liens contextuels par le transfert des 'ancres' cliquables à du texte normal.

❖ L'annotation adaptative

L'idée est d'augmenter les liens avec quelques formes de commentaires qui peuvent renseigner l'utilisateur sur l'état actuel des nœuds derrière les liens annotés. Ces annotations peuvent être fournies sous forme textuelle ou visuelle en utilisant, par exemple, différentes icônes, couleurs ou tailles de caractères. Le lien d'annotation est connu comme une technique efficace de support de navigation dans les hypermédias.

❖ L'adaptation des cartes

La technique d'adaptation des cartes comprend diverses façons d'adaptation de la forme des cartes hypermédia locales et globales présentées à l'utilisateur. Des techniques comme le guidage direct, le masquage des liens et l'annotation peuvent être utilisées aussi pour adapter les cartes hypermédia, mais toutes ces techniques ne changent pas la forme ou la structure des cartes.

5.6. Conclusion

Dans ce chapitre, nous nous sommes intéressés au concept d'adaptation qui est né avec les nouveaux systèmes d'apprentissage à distance et qui sont centrés sur le document pédagogique. Nous avons défini les deux termes : adaptation et adaptabilité en donnant la différence entre ces deux derniers termes et le troisième terme d'adaptativité. Nous avons parlé de la flexibilité et de l'expérience de l'utilisateur dans un système éducatif en ligne. Puis, nous avons donné les différents types de systèmes intégrant l'adaptation.

Ensuite, nous avons expliqué les critères d'adaptation. Avant de détailler les différentes techniques et méthodes d'adaptation, nous avons les précédé par les paramètres d'adaptation, ses modèles et ses éléments. Les techniques d'adaptation sont divisées en deux grandes catégories : la présentation adaptative et le support de navigation adaptative. Dans chacune des deux, nous avons détaillé les méthodes et les techniques.

Ce chapitre est le dernier dans la première partie, la deuxième partie se concentre sur notre projet : 'PrOALDoF-Web'.

PARTIE – II : NOTRE PROJET

PrOALDoF-Web

Contenu

Chapitre I : *Présentation du PrOALDoF-Web*

Chapitre II : *La Recherche dans PrOALDoF-Web*

Chapitre III : *La Fusion et l'Adaptation dans PrOALDoF-Web*

Chapitre VI : *Evaluation du PrOALDoF-Web*

6. Présentation du PrOALDoF-Web

6.1. Introduction

La technologie a créé le cyberspace et décomposé les barrières nationales et internationales. Aujourd'hui, n'importe qui possède un ordinateur, un modem et un fournisseur d'accès qui peut permet de connecter aux nouveaux moyens et profiter de ses nombreux avantages. Tout le monde peut enseigner et n'importe qui peut apprendre quoi, quand et où il veut. Actuellement, la révolution technologique rapide a changé le monde complètement, surtout au niveau des connaissances et de transfert de l'information à l'intérieur et à l'extérieur de l'université. Ce transfert d'information donne ainsi un nouveau type d'enseignement supérieur fondamentalement différent du précédent (Brusilovsky, 1999).

L'éducation en ligne comme un outil d'apprentissage à distance est considérée comme une méthode d'enseignement qui permet à l'enseignant et à l'apprenant d'être séparés les uns des autres dans le temps et dans l'endroit pendant qu'ils sont reliés par une technologie appropriée (Kolb, 1984). Dans l'enseignement en ligne, la séparation du temps et d'endroit est assurée par les réseaux des TIC 'Technologies de l'Information et de la Communication', notamment l'Internet et le WWW 'World Wide Web'.

E-Learning est utile dans l'éducation, le commerce et tous les types d'apprentissage. Grâce à cette méthode, l'argent et le temps peuvent être réduits. En plus, en produisant des résultats mesurables et très satisfaisant [(Amaral et Leal, 2006), (Arthurs, 2007), (Bostrom, Olfman et Sein, 1990), et (Brusilovsky, 2004)]. E-Learning est principalement le réseau du transfert des compétences et des connaissances dans la plateforme de formation en utilisant des applications électroniques et un processus d'apprentissage [(Brooks, Kettel et Hansen, 2005), (Rodica et Anca, 2008)]. Avec le développement rapide de la technologie numérique, l'information est devenue un domaine de recherche très important à travers lequel les chercheurs ouvrent des portes sur plusieurs domaines la recherche, telles que l'indexation, la sauvegarde et la représentation de l'information (Senan, 2008).

L'information est responsable pour chaque préparation des cours d'une formation en ligne, mais elle exige aussi des méthodes, en particulier la sélection des informations les plus pertinentes (Chaoui et Laskri, 2011b). L'information est disponible sous plusieurs formats sur le Web, en particulier sous format textuel souple pour toute modification et mise à jour qui peuvent être réalisées par des auteurs et des apprenants ou internautes.

Cette possibilité d'ouverture et de liberté rend la tâche d'adaptation des informations un processus très difficile qui peut être non terminé [(Canales et *al*, 2007), (García Barrios, 2006), (Monova-Zheleva, 2005), (Rosmalen et *al*, 2006) et (Ruiz et *al*, 2008)]. En outre, la quantité du matériel d'apprentissage sur Internet a connu une croissance rapide au cours des dernières décennies.

Par conséquent, les consommateurs de l'information sont mis au défi de choisir les bons équipements. Dans les systèmes d'apprentissage en ligne, la plupart des approches ont conduit une confusion chez les apprenants. Pour cela, l'apprentissage adaptatif a gagné beaucoup d'attention dans ce domaine [(Chaoui et Laskri, 2011a), (Monova-Zheleva, 2005), (Papanikolaou et *al*, 2002), (Rey-Lopez et *al*, 2009), (Sampson, Karagiannidis et Cardinali, 2002), (Wang, Wang et Huang, 2008), (Yang et Wu, 2009) et (Yessad et *al*, 2008)].

Ce chapitre a comme but la présentation du projet PrOALDoF-Web en commençant par une motivation, puis une description du projet et ses objectifs, par la suite nous détaillons l'architecture générale du projet avec ses différentes parties. Enfin, nous donnons l'algorithme de l'approche proposée avec le principe du fonctionnement de tout le projet.

6.2. Motivation

Avec l'avènement de l'Internet, la très grande masse d'informations devenue disponible nécessite aujourd'hui de consacrer une partie considérable de notre temps à l'extraction de l'information pertinente. L'enseignant ou bien le constructeur du cours équipé seulement d'outils de recherche d'information sur les réseaux tels que le WWW, ne peut pas faire face au flux d'information générée. Au lieu de laisser l'enseignant dépenser son temps à chercher l'information dont il a besoin, la tendance actuelle est de concevoir des mécanismes qui permettent de lui faciliter la tâche en lui faisant parvenir l'information qui l'intéresse.

Et avec l'avènement des ontologies, les enseignants créent des bases de cours (domaine d'apprentissage) en se basant sur plusieurs ontologies, par exemple une ontologie de domaine décrivant les concepts du domaine de la formation ; une autre pédagogique décrivant l'ensemble des notions utiles à une formation particulière ; et nous pouvons intégrer plusieurs ontologies dont le but est de construire le cours dynamiquement via les concepts existants, toutes les ressources disponibles dans la formation et bien sur le comportement de l'utilisateur (son état actuel, ses connaissances, son état psychologique).

Une telle démarche manuelle pour trouver les briques d'informations nécessaires dans la formation est répétitive et couteuse en temps. Du fait que, les informations pertinentes sont très éparpillées dans le Web, ainsi les documents contenant ces informations sont nombreux et peuvent avoir une grande taille en terme nombre de mots, ce qui demande un effort pour leurs lectures et leurs compréhensions. De plus, les index des pages Web des moteurs de recherche sont en constante mise à jour.

Notre problématique de recherche consiste à améliorer la construction manuelle vers une construction automatique. Donc, dans notre projet, nous avons proposé une approche pour créer un domaine d'apprentissage sans aucune intervention des concepteurs (enseignants).

L'approche commence par : trouver les documents les plus pertinents qui se trouvent sur le Web, les télécharger, les filtrés en se basant sur une ontologie de domaine et sur les règles sémantiques, enfin extraire les parties les plus pertinentes en appliquant la méthode de fusion proposée afin de créer le domaine d'apprentissage de l'approche proposée. Par la suite, une adaptation basée fusion sera effectuée pour créer les cours dynamiquement à partir du domaine d'apprentissage créé dont le but est d'obtenir un cours adaptatif au profil de l'apprenant.

Notre objectif à travers cette approche proposée est de réduire l'immense espace du Web qui contient des milliards de pages Web. Cet espace devient personnalisé avec l'adaptation directe des cours aux profils des apprenants afin de répondre à leurs besoins et afin d'augmenter leurs satisfactions en fournissant une bonne formation extensible à tout changement et à toute mise à jour, mais avec des ressources éducatives fiables.

6.3. Description du PrOALDoF-Web

PrOALDoF-Web est un projet de recherche qui consiste à réaliser un système d'enseignement en ligne dont le contenu pédagogique est adaptatif au profil de l'apprenant et dont la source d'information disponible sur le Web est bien entendu extraite et réorganisée selon un canevas prédéfini du domaine d'enseignement afin d'enrichir convenablement les connaissances de l'apprenant selon ses aptitudes et selon son niveau. PrOALDoF-Web est une abréviation en anglais "**Proposition and Organization of an Adaptive Learning Domain based on Fusion from the Web**". Donc, PrOALDoF-Web propose une nouvelle approche et organise un domaine d'apprentissage adaptatif au profil de l'apprenant en se basant sur la méthode de fusion des ressources disponibles dans le Web.

Pour répondre aux objectifs de la thèse, nous donnons notre approche :

- L'utilisation de l'API Google pour créer un corpus documentaire qui joue le rôle d'un support physique de test de l'approche proposée. Nous pouvons rajouter manuellement d'autres documents, car il y a une limite dans les requêtes et le nombre de résultats par jour du moteur de recherche Google ;
- La proposition d'une méthode de recherche des documents les plus pertinents (plus précisément, les morceaux documentaires les plus pertinents) en se basant sur une ontologie de domaine ;
- La proposition d'une méthode de fusion de morceaux de texte extraits à partir du corpus ;
- L'adaptation du contenu pédagogique en utilisant la méthode de fusion et une ontologie pédagogique.

PrOALDoF-Web est composé de trois parties : 1) Partie Recherche, 2) Partie Fusion, et 3) Partie Adaptation. Les trois parties peuvent travailler en simultané ou individuellement, mais à la fin du processus de traitement, les trois parties deviennent complémentaires.

A travers PrOALDoF-Web, nous développons une nouvelle approche dans les systèmes éducatifs en ligne adaptatifs, qui peut être intégrée à n'importe quel LMS. Le contenu d'apprentissage peut être réutilisé et partagé entre les différentes plates-formes. La source d'information de contenu est le Web. La méthode de recherche et de filtrage est basée sur une ontologie de domaine et sur les règles sémantiques. L'approche d'adaptation est basée sur la méthode de fusion en utilisant l'ontologie de domaine décrivant les concepts du cours et l'ontologie pédagogique décrivant la technique d'adaptation. Ce qui est nouveau dans notre travail est la fusion de plusieurs sous fragments des ressources Web jugées pertinentes. En conséquence, l'amélioration de la qualité du contenu de la formation et l'obtention d'un système éducatif adaptatif, fiable, riche et dynamique ; dynamique est dans le sens d'enrichissement et de mise à jour.

6.3.1. Partie Recherche

Elle sert à rechercher sur le Web via une requête formulée d'une façon automatique, puis à filtrer les résultats (documents) obtenus pour créer une base documentaire. Cette partie nécessite une connexion internet, donc, elle est sous le mode en ligne. La base documentaire va être stockée au niveau du serveur de la formation (Voir Chapitre II, Partie II en page 106).

6.3.2. Partie Fusion

Dans cette partie, un processus d'extraction se lance pour trouver les sous fragments les plus pertinents, puis les fusionner pour construire de nouveaux fragments, ces derniers vont être sauvegardés dans une **Base de Nouveaux Fragments 'BNF' (New Fragments DataBase 'NFDB')**. Cette base présente un domaine d'apprentissage 'DA' (Learning Domain 'LD'). 'DA' dernier offre la possibilité de création des cours qui doivent contenir toutes les briques d'informations nécessaires dans la formation. 'DA' est en constante mise à jour selon le choix des auteurs de la formation, qui vont contrôler les briques d'informations sauvegardées dans la base 'BNF' (Voir Chapitre III, Partie II en page 132).

6.3.3. Partie Adaptation

Dans cette partie, la base 'BNF' joue un rôle très important dans la création des cours adaptatifs aux profils des apprenants en se basant sur l'approche de Fusion proposée. Cette dernière, organise le cours selon les méthodes et les techniques de la présentation adaptative. La Fusion crée un domaine d'apprentissage adaptatif 'DAA' (Adaptive Learning Domain 'ALD'). Ce nouveau domaine est particulier pour chaque utilisateur dans la formation. 'DAA' est en constant enrichissement en relation avec le profil de l'apprenant (Voir Chapitre III, Partie II en page 132).

Les deux parties Fusion et Adaptation, n'obligent pas une connexion internet si les utilisateurs sont dans le corpus, mais surtout une connexion au réseau local au serveur de la formation. Elles peuvent fonctionner sous le mode off ligne. La connexion est souhaitable en cas d'absence d'un sous fragment dans la base 'BNF', car dans ce cas, la partie Recherche va être relancée pour enrichir la base avec les ressources manquantes.

En effet, la connexion via le réseau local va diminuer l'accès extérieur (accès internet). Parce que, la création du corpus documentaire se déclenche dès la première utilisation, cette dernière est lancée par un auteur de la formation qui va contrôler son contenu après le lancement du processus de fusion afin de créer un domaine d'apprentissage contenant tous les fragments nécessaires dans la formation. La mise à jour du corpus, se déclenche dès qu'une ressource manquante dans un 'DAA' d'un utilisateur est détectée après avoir utilisé le système de formation. Le système peut ne localise aucune ressource Web adéquate à la ressource manquante. Dans ce cas, il y a deux problèmes qui peuvent être survenus.

Le premier, la recherche n'est pas fructueuse à cause d'une mauvaise formulation de la requête de la recherche Web. L'auteur doit signaler ce problème au concepteur de l'ontologie de domaine. Le concepteur va modifier les annotations du concept ou sous concept en cause, ou bien l'enrichir avec de nouvelles annotations.

Le deuxième, l'extraction des sous fragments les plus pertinents n'est pas fructueuse à cause d'une erreur au niveau des règles sémantiques impliquées dans la phase d'analyse pour les comparées avec les sous fragments des ressources Web. A cet état, l'auteur doit signaler le problème à l'expert qui a créé les règles sémantiques. L'expert va régler le problème en corrigeant la règle ou l'ensemble des règles ayant une erreur.

6.4. Architecture de l'approche proposée 'PrOALDoF-Web'

Dans cette partie, nous allons détailler notre approche proposée 'PrOALDoF-Web' en schématisant l'architecture générale de l'approche proposée (Voir Figure 6 ci-dessous). Le but est de construire les différentes parties de notre système.

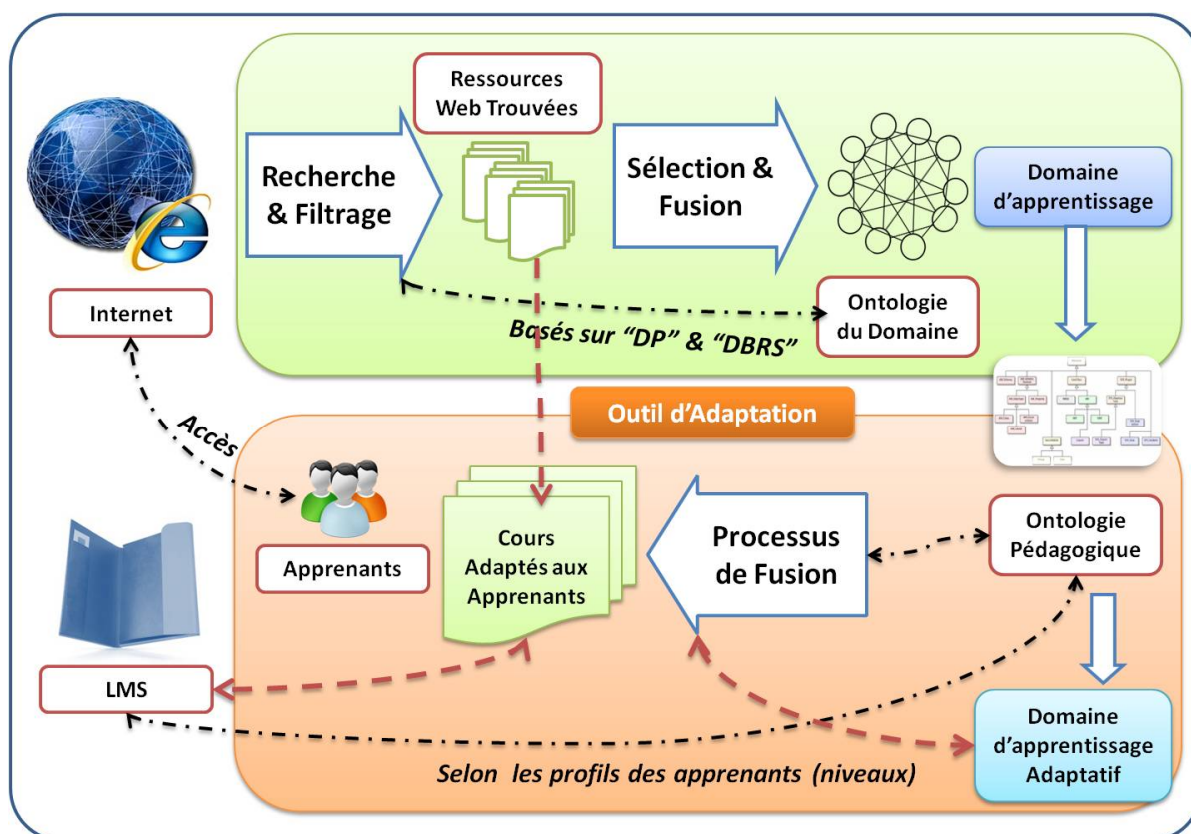


Figure 6 L'architecture générale de PrOALDoF-Web

L'approche propose un nouvel outil performant dans la recherche et le filtrage des ressources Web basé sur le degré de pertinence 'DP' (Degree of Relevance 'DR') et sur la Distance Basée sur des Règles Sémantiques 'DBRS' (Distance Based on Semantic Rules 'DBSR'). Les deux méthodes utilisent l'ontologie de domaine.

Nous organisons notre cours selon la hiérarchie présente dans l'ontologie de domaine pour construire un domaine d'apprentissage 'DA'. Ensuite, nous utilisons l'ontologie pédagogique pour construire un domaine d'apprentissage adaptatif 'DAA'. L'ontologie pédagogique dans notre travail contient les niveaux des apprenants et la stratégie d'adaptation basée sur la méthode de fusion. 'DAA' présente une base de données contenant les ressources Web les plus pertinentes. Avec l'application de l'approche de fusion, le contenu du cours peut avoir une très bonne qualité avec la possibilité de réutilisation des ressources Web déjà extraites et sauvegardées dans la Base des Nouveaux Fragments 'BNF'. L'architecture générale de PrOALDoF-Web facilite la compréhension du processus de fonctionnement afin d'obtenir des cours adaptatifs aux apprenants en cherchant leurs contenus via le Web.

Pour créer un cours adaptatif dans le système PrOALDoF-Web, nous avons cinq étapes : (1) : Recherche sur le Web, (2) : Filtrage et Sélection, (3) : Fusion des sous fragments, (4) : Construction du domaine d'apprentissage et (5) : Adaptation aux profils des apprenants. Pour mieux comprendre l'approche proposée, nous réalisons l'algorithme appelé 'PrOALDoF-Web', et nous donnons une description textuelle de l'algorithme (Voir Figure 7 en page 102), ainsi, une description algorithmique (en page 103). Nous détaillons les principales étapes dans l'approche proposée et surtout ce qui est nouveau dans notre projet par rapport aux projets similaires. Douze étapes sont mentionnées dans la description textuelle.

L'approche est caractérisée par une méthode qui filtre les ressources du Web en appliquant des méthodes statistiques et sémantiques basées sur les concepts de l'ontologie et sur une liste de mots-clés pour chaque concept, afin de définir la liste des documents les plus pertinents pour chaque partie du cours. Après cela, nous appliquons la méthode sémantique (selon les règles sémantiques) pour chaque fragment d'un document entre ses mots et les termes définis dans chaque concept de l'ontologie. En conséquence, les deux méthodes sont complémentaires pour affecter un ensemble de fragments pour chaque concept qui présente une partie du cours juste après l'extraction des fragments les plus pertinents. Ces fragments sont stockés dans des bases de données, c'est-à-dire, nos deux méthodes fonctionnent dès qu'un nouveau document Web est détecté pertinent, de façon à minimiser le temps.

Avec l'approche de fusion, les cours peuvent être créés à partir des différents fragments qui sont extraits à partir d'une ou plusieurs ressources Web. L'Adaptation du contenu du cours est basée sur la fusion. Le contenu du cours peut être modifié en fonction du niveau des apprenants. Lorsque le niveau augmente, l'enrichissement des cours augmente aussi.

Algorithme PrOALDoF-Web ;

Début

- 1- Rechercher des ressources dans le Web via Google API, avec la formulation automatique des requêtes de recherche en se basant sur l'ontologie de domaine ; chaque concept a un ensemble de mots clés ;
- 2- Filtrage basé ontologie de domaine. Dans un premier temps, il y a le calcul du degré de pertinence 'DP' pour réduire l'espace de filtrage. Dans un deuxième temps, il y a le calcul de la distance entre les sous fragments pour chaque fragment d'une ressource Web trouvée, la distance est basée sur les règles sémantiques 'DBRS' ;
- 3- Sélection des ressources les plus pertinentes ;
- 4- Sauvegarde des fragments dans la base des nouveaux fragments 'BNF' ;
- 5- Fusion du contenu des sous fragments en se basant sur la distance entre ces derniers et sur les règles sémantiques 'DBRS' ;
- 6- Associer chaque nouveau fragment à une partie du cours après la fin de l'étape de fusion ;
- 7- La Sauvegarde des nouveaux fragments dans la base 'BNF' ;
- 8- Relier 'BNF' avec l'ontologie pédagogique pour créer un domaine d'apprentissage 'DA' ;
- 9- Relier 'DA' aux apprenants ;
- 10- Adapter les cours aux apprenants pour créer un domaine d'apprentissage adaptatif 'DAA' en se basant sur la fusion à partir le Web ;
- 11- Enrichir convenablement la base 'BNF' à partir du Web ;
- 12- Les ressources du domaine 'DAA' peuvent être réutilisée directement à partir la base 'BNF' ;

Fin.

Figure 7 Description textuelle de l'algorithme PrOALDoF-Web

6.5. Description algorithmique du fonctionnement de PrOALDoF-Web

Cette description présente notre proposition sous format algorithmique comme suit :

Algorithme PrOALDoF-Web ;

Procédure OntoToExcel(Ontology, BNF) ;

Fonction DP(Concept, Fragment) ;

Fonction DBRS(RS, Sous_Fragment) ;

Fonction Requete_de_Recherche(ontologie) ;

Procédure Parcourir_Documents_Web(Requete) ;

Procédure Télécharger(Document) ;

Procédure Parcourir_Documents (Document) ;

Procédure Sauvegarder(Degré, Base) ;

Procédure Parcourir_Sous_Fragments (Sous_Fragment);

Procédure Sauvegarder(Distance, Base) ;

Procédure Fusionner(concept, BNF) ;

Procédure Relier(BNF, ontologie) ;

Procédure Adapter(apprenant, DAA, BNF) ;

Procédure Recherche_Web() ;

 Début

 Parcourir_Documents_Web(Requete_de_Recherche(ontologie)) ; //“trouver
 //les documents pertinents”

 Si Document est Pertinent Alors

 Télécharger(Document) ;

 Fin

Procédure Filtrage() ;

 Début

 Parcourir_Documents (Document) ; //“calcul de tous les degrés de pertinence
 //des concepts de l’ontologie par rapport aux fragments des documents
 //pour associer chacun des derniers à un seul concept”

 Sauvegarder(DP(Fragment_Web), BNF) ; //“ordre des DP du même concept,
 //BNF : Base des Nouveaux Fragments”

 Fin

Procédure Fusion() ;

Début

Parcourir_Sous_Fragments(Sous_Fragments) ; //“calcul de toutes les
//distances entre les termes des phrases (sous fragments) d’un fragment du
//document en se basant sur les règles sémantiques du concept pour
//associer chaque sous fragment à une seule règle sémantique”

Sauvegarder(DBRS(RS, Sous_Fragment), BNF) ; //“ordre des DBRS de la
//même règle sémantique d’un concept”

Fusionner(concept, BNF) ; //“fusion des sous fragments avec les distances
//‘DBRS’ les plus grandes et sauvegarder les nouveaux fragments dans la
//base ‘BNF’”

Fin

Procédure Adaptation() ;

Début

Relier(BNF, ontologie) ; //“mise à jour de la base ‘BNF’ en donnant pour
//chaque règle sémantique un seul niveau selon la stratégie d’adaptation
//dans l’ontologie pédagogique”

Adapter(apprenant, DAA, BNF) ; //“adapter les cours aux apprenants en
//choisissant les ressources de la formation à partir la base ‘BNF’ afin de
//créer l’espace réservé ‘personnalisé’ à chaque apprenant qui représente
//un domaine d’apprentissage adaptatif ‘DAA’, cette méthode est basée sur
//l’approche de fusion”

Fin

// Programme Principal

Début

Procédure OntoToExcel(Ontology, BNF) ;

Recherche_Web() ;

Filtrage() ;

Fusion() ;

Adaptation() ;

Fin.

6.6. Conclusion

Nous avons présenté via ce chapitre notre projet PrOALDoF-Web en commençant par une motivation, puis une description du projet et ses objectifs, par la suite nous avons détaillé l'architecture générale du projet avec ses différentes parties. Enfin, nous avons donné l'algorithme de l'approche proposée avec le principe du fonctionnement de tout le projet. Nous parlons sur la recherche et la représentation des connaissances dans PrOALDoF-Web dans le chapitre suivant. Cette partie du travail concerne la première tâche de notre approche proposée.

7. La Recherche et la représentation des connaissances dans PrOALDoF-Web

7.1. Introduction

Ce chapitre présente les différentes étapes de la recherche dans le projet PrOALDoF-Web en commençant par une définition du terme recherche documentaire. Ensuite, nous citons les types des documents. Puis, nous parlons sur les sources documentaires. Les étapes de notre recherche sont détaillées dans le point suivant en commençons par la formulation automatique de la requête de recherche, trouver les informations pertinentes et enfin analyser le contenu des informations trouvées. Puis, une représentation des connaissances ‘Ontologie de domaine’ est définie. L’élément suivant, est bien entendu le domaine d’application qui est dans notre cas le squelette humain dans lequel nous avons créé notre ontologie de domaine et notre ontologie pédagogique. La méthodologie ontologique Uschold et King’s est celle utilisée pour créer nos ontologies qui sont intégrées via l’API JENA. Enfin, nous détaillons notre méthode de filtrage pour extraire les informations les plus pertinentes.

7.2. La Recherche documentaire

La recherche documentaire est l'utilisation de sources extérieures des documents pour soutenir le point de vue ou l’argument d'un travail universitaire. Le processus de recherche documentaire implique souvent une ou toutes les parties de la conceptualisation, l'utilisation et l'évaluation des documents. L'analyse des documents dans la recherche documentaire serait soit une analyse quantitative ou qualitative (ou les deux).

La principale question entourant les types de documents est de notre capacité à les utiliser comme sources fiables de données dans le monde académique (ou autre) qui doit être considéré par tous les chercheurs qui utilisent des documents dans leurs recherches. L’insuffisance des sources disponibles jusqu’à présent signifie que ce recueil sera inestimable pour les chercheurs.

Devant la masse des informations textuelles disponibles, notamment sur le Web, la recherche documentaire nécessite un traitement automatique pour une analyse pertinente de ces données. Cela s’avère particulièrement utile pour découvrir les informations stratégiques afin d’anticiper et de prendre des décisions sur le programme (le cours à construire par la suite dans les prochaines étapes du projet) d’un apprenant.

7.3. Les Type de documents

Il y a plusieurs exemples de type de documents qui comprennent des publications gouvernementales, des journaux, des certificats, des publications de recensement, romans, films et vidéos, peintures, photographies personnelles, agendas et d'innombrables autres écrits, visuels et sources picturales dans deux types différents de format : imprimé ou électronique. Il existe deux types de documents qui peuvent être utilisés dans l'étude documentaire, les documents à savoir primaires et secondaires. Les documents primaires font référence à des témoignages produits par les gens qui ont vécu un événement particulier ou le comportement qu'ils veulent étudier. D'autre part, les documents secondaires sont des documents produits par des gens qui n'étaient pas présents sur les lieux, mais qui ont reçu des témoignages pour compiler les documents, ou qui ont lu les témoignages.

7.4. Les sources documentaires

Un document est un artefact qui a comme caractéristique essentielle un texte inscrit. Autrement dit, un document est un texte écrit. Les documents sont produits par des individus et des groupes dans le cadre de leurs pratiques quotidiens en enseignement ou en formation, ou en visant exclusivement dans d'autres cas leurs propres besoins pratiques immédiats en recherche ou autres.

La production des documents suit un but bien précis et ils sont fondés sur certaines hypothèses et ils sont présentés d'une certaine manière ou style. Un chercheur par exemple, doit être pleinement conscient de l'origine, le but et le public original des documents. Il faut noter que les documents ne sont pas délibérément produits à des fins de recherche, mais les objets d'origine naturels avec une existence concrète ou semi permanente qui nous dirigent indirectement vers le monde social des gens qui les ont créés. Un document, contrairement à un discours, peut avoir une existence indépendante au-delà de l'écrivain et au-delà du contexte de sa production. Les documents peuvent être de nature publique, privée ou personnelle.

La liste des sources de documents publique comprend les publications gouvernementales telles que les lois du Parlement, des énoncés de politique, des rapports de recensement, les bulletins statistiques, les rapports des commissions d'enquête, des paroles des ministres ou des rapports annuels des ministères, rapports des assemblées, ...etc.

Les documents privés proviennent souvent des organisations de la société civile telles que les entreprises du secteur privé, les syndicats et les organisations non gouvernementales, ainsi bien sûr des particuliers. Ils comprennent les rapports des réunions, les résolutions du conseil d'administration, publicités, factures, dossiers du personnel, manuels de formation, et autres rapports annuels, ...etc.

Enfin, la liste des documents personnels comprend des livres, carnets d'adresses, des dossiers médicaux, agendas, lettres personnelles, ...etc.

Pour permettre aux concepteurs d'extraire les informations les plus pertinentes, PrOALDoF-Web propose une nouvelle méthode de recherche en se basant sur le moteur de recherche 'Google'.

7.5. Etapes de la recherche

Nous avons trois étapes (1) : la formulation automatique de la requête de recherche, (2) : trouver les informations pertinentes et (3) : analyser le contenu des informations trouvées (Voir Figure 8 ci-dessous pour plus de détails).

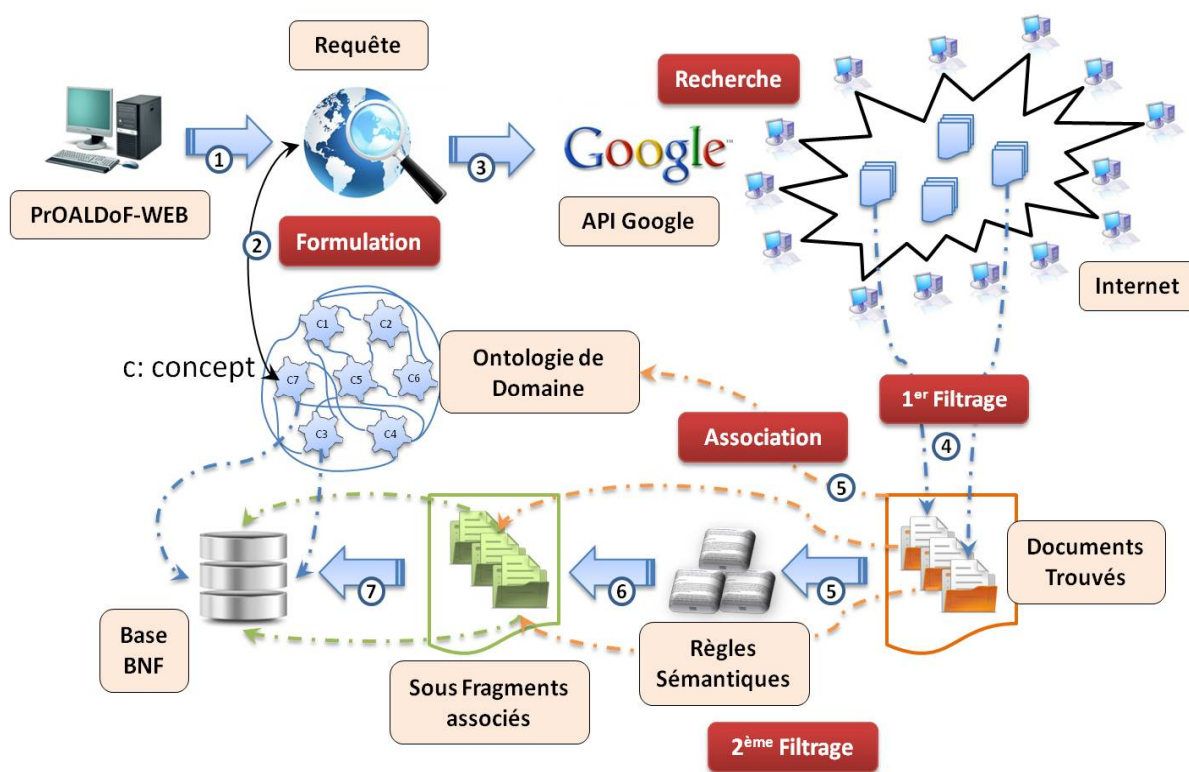


Figure 8 Etapes de la Recherche dans PrOALDoF-Web

7.5.1. La formulation automatique de la requête de recherche

Dans cette étape le système extrait l'ensemble des mots-clés regroupés dans la zone des annotations du concept principal dans l'ontologie de domaine. Le concept étant spécifique d'un domaine de recherche (squelette humain). Les mots-clés sont définis lors de l'analyse du document par le moteur de recherche (Google). Ils décrivent le plus précisément possible le contenu du document. L'utilisation des annotations ontologiques définissantes par un ou plusieurs experts, permet d'avoir des termes univoques pour être sûr d'obtenir le maximum de pages Web (adresses URL) par rapport à la question posée.

Chaque annotation est validée par l'expert pour respecter l'orthographe. Ainsi, pour associer les bons opérateurs booléens (ET, OU, SAUF) aux termes de recherche, afin de réaliser une équation de recherche efficace.

Il est important de bien formuler la requête de recherche pour réduire les silences et les bruits.

7.5.2. Trouver les informations pertinentes

La recherche d'informations consiste à trouver, dans une base de documents de taille importante, les documents pertinents correspondant à des requêtes par mots-clés formulées dans l'étape précédente. La recherche correspond à la partie requêtes du moteur, qui restitue les résultats. Un algorithme est appliqué afin de pouvoir présenter les résultats des recherches selon un ordre de pertinence. L'algorithme tient généralement compte du contexte du mot clé (titre, paragraphe, hyperlien...) et de la ressource (ressources liées, popularité du site...). La recherche est généralement effectuée en indexant préalablement tous les documents de la base selon les mots qu'ils contiennent ; la recherche consiste à trouver, le plus rapidement possible, les documents ayant des mots communs avec la requête formulée.

Dans PrOALDoF-Web, l'objectif principal de cette étape, est de fournir un ensemble de documents pertinents. Cet ensemble définit ce que l'on appelle, dans la suite de ce document, un cours, thème ou une catégorie. Pour un cours donné, une hiérarchie du contenu est définie dans l'ontologie de domaine. Comme il y a une hiérarchie contenant des sous titres (sous concepts dans l'ontologie), le résultat voulu n'est pas les documents pertinents par rapport seulement au concept principal du sujet de recherche, mais aussi, par rapport aux sous concepts de ce dernier.

Le choix des mots-clés pour la génération automatique de la requête est approprié à l'expert du domaine pris en charge à la création de l'ontologie. Dans notre cas, l'ensemble des mots-clés sont bien écrits dans l'espace réservé par Protégé. Le système charge les mots-clés depuis l'ontologie, et génère les requêtes de recherche pour trouver les documents demandés. Le but donc de cette étape est de chercher sur le Web via les requêtes ; bien formulées automatiquement par le système ; les documents pertinents par rapport aux questions posées.

7.5.2.1. Quel type d'information

Nous intéressons dans notre projet par les informations textuelles non structurées ou semi structurées. Dans notre projet, les deux types de document Word et PDF sont pris en charge. Par expérience, les bons documents éducatifs qui se trouvent sur le Web sont des documents de type (PDF et Word).

Dans une recherche de type "information", il s'agit plutôt de retrouver des éléments d'information sur un thème donné : des statistiques, des données biographiques, une synthèse sur le sujet, ...etc. Ainsi, retrouver le texte d'une œuvre (source primaire) ou d'un ouvrage ou d'un article de critique (source secondaire).

Les résultats de la recherche sont les URL des deux documents (PDF et Word). Ce premier filtrage est assuré par le moteur de recherche Google en utilisant l'API Google. Ce qui nous donne une nouvelle définition de la nature des informations dans PrOALDoF-Web. Les résultats trouvés sont de nature différente, elles peuvent être des articles, des ouvrages, des synthèses ...etc.

7.5.2.2. Juger la pertinence des informations

L'objet principal du texte de la recherche est celui de la requête générée précédemment. Le moteur de recherche trouve de nombreux mots synonymes au sujet de recherche, car le moteur de recherche indexe les pages aussi bien par rapport à ses mots (contenu) que par rapport aux mots liés au sujet de chaque page indexée. Ce qui nous permet d'élargir le volume des résultats de la recherche.

Les moteurs de recherche tentent de définir de manière plus précise le contenu et la qualité d'un document en introduisant des critères de pertinence par rapport à chaque mot contenu dans ce document. Cette pertinence est définie en utilisant les critères suivants :

- Les mots rencontrés et leurs fréquences d'apparition dans le contenu de chaque document.
- Les mots sont dans : un titre, l'adresse du document, les premiers paragraphes.
- Les mots sont en gras, ou dans un autre enrichissement typographique.
- Les mots sont dans les Méta tags du document.
- Les mots demandés sont rapprochés les uns des autres.
- Les mots sont dans le même ordre que ceux de la requête.
- L'indice de popularité de la page est élevé. La base de données d'un moteur possède quelques millions de pages. Il est possible de calculer combien d'autres pages de la base de données pointent vers la page analysée (possèdent un lien hypertexte vers elle). Google appelle cet indice de Popularité 'Page Rank'.

En utilisant le moteur de recherche Google, nous assurons une haute pertinence surtout avec les premiers résultats trouvés. Donc, effectuer la recherche des documents avec l'API Google donne une forte possibilité de trouver les résultats souhaités.

7.5.2.3.L'API Google et la recherche

L'API Google est un kit de développement logiciel disponible librement, qui permet de créer de nouvelles applications en utilisant directement la base de données des pages indexées par Google, par le biais d'un service web.

Donc, nous utilisons cette API pour réaliser notre module de recherche automatique via l'ontologie de domaine. Cette API effectue une communication avec le serveur de Google par l'intermédiaire d'un protocole SOAP 'Simple Object Access Protocol', pour le traitement des informations entrées par le module de recherche.

Nous utilisons l'API Google automatiquement dans notre système au lieu d'utiliser la recherche manuelle à travers le moteur de recherche Google. Cette API fournit la liste des résultats représentant les URL obtenues à partir des requêtes formulées automatiquement à partir des mots-clés et des annotations qui se trouvent dans l'ontologie de domaine. L'API Google permet de filtrer les résultats par types des documents (PDF, Word, ...etc.) qui sont donnés par la commande : 'filetype'. Par l'application de l'API Google dans notre projet, chaque partie du cours génère une requête. Une telle recherche manuelle donne les mêmes résultats que notre méthode, car nous utilisons le même moteur de recherche 'Google' (Voir Figure 9 en page 112)

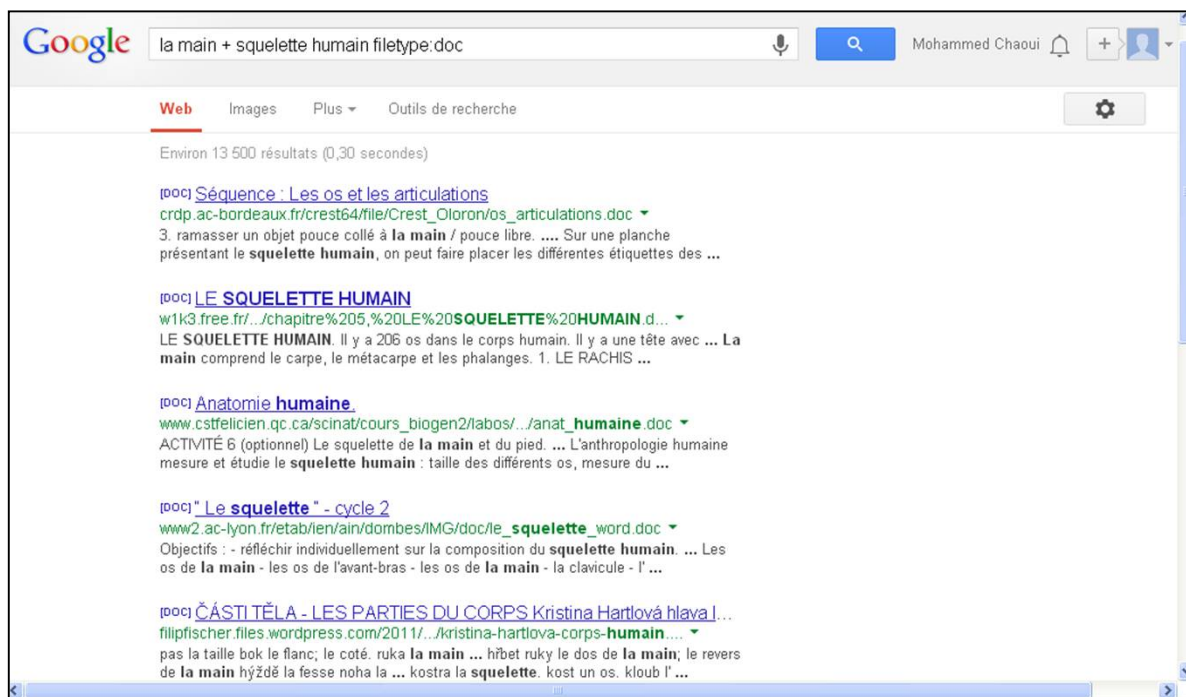


Figure 9 Exemple de recherche personnalisée (Word) dans Google

7.5.2.3.1. Un exemple de recherche par cette API

- L'ontologie de domaine : Squelette Humain ;
 - Cours : Squelette humain ;
 - Partie du cours : La main ;
 - Requête : La main + Squelette Humain filetype:doc.
- La main + Squelette Humain : cette combinaison présente la requête de recherche.
 - filetype:doc : spécifie le critère de filtrage des documents dans cet exemple de recherche.

Le résultat est une liste d'URL et les documents sont de type Word (doc) (Voir Figure 9 ci-dessus).

Que ce soit les documents sont de type Word ou PDF, la méthode fonctionne de la même manière, sauf un petit changement dans la requête de recherche. La requête de l'exemple précédent devient comme suit : La main + Squelette Humain filetype:pdf. Cette requête fait aussi un appel à l'API Google pour avoir une liste d'URL et les documents sont de type PDF (Voir Figure 10 en page 113).

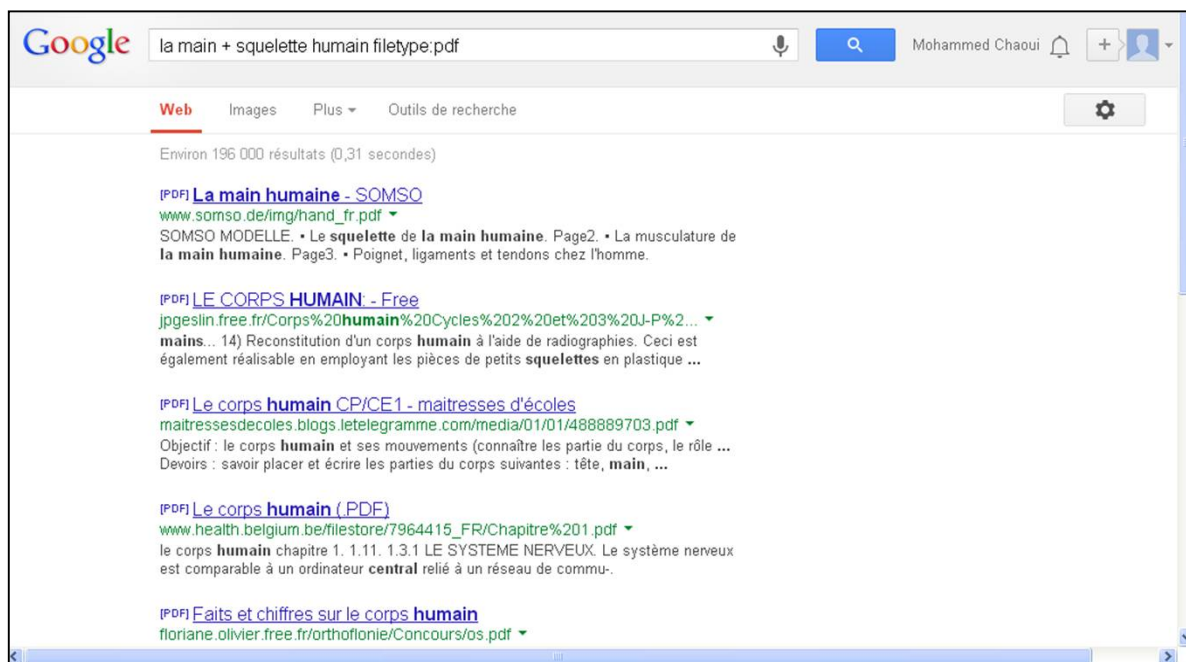


Figure 10 Exemple de recherche personnalisée (PDF) dans Google

Pour toute requête dans PrOALDoF-Web, la recherche se lance deux fois pour les deux types de document (Word et PDF). En pratique, nous ne pouvons pas lancer la recherche des deux types en une seule requête, cela est autorisé uniquement si nous effectuons une requête avec l'opérateur booléen 'OR'. Dans ce cas, nous pouvons utiliser des synonymes par exemple dans la recherche ('Squelette humain' or 'Système squelettique humain').

A la fin de cette étape, PrOALDoF-Web est pris en charge pour sauvegarder tous les documents dans la base locale.

7.5.2.4.Sauvegarde des informations

Une telle méthode manuelle, consiste à cliquer sur un lien. Pour notre cas il faut lancer le téléchargement pour tous les documents automatiquement. C'est-à-dire, si la liste des URL contient par exemple 10 URL, il faut répéter l'action du lancement 10 fois, mais ça devient purement manuel. Pour cela nous avons pensé aux logiciels de téléchargement.

Il existe pas mal d'outils de téléchargement, qui ont comme objectif la sauvegarde des fichiers de types différents (Audio, documents, Vidéo, ...etc.). Mais, l'implémentation du code source d'un parmi ces logiciels dans notre système est impossible car ils ne sont pas Open Source.

Pour cela, notre stratégie est d'implémenter un module de téléchargement à l'intérieur du système. Nous avons les URL obtenus à partir de l'étape de la recherche. Donc, lors du traitement de la recherche, un autre module de téléchargement sera lancé pour sauvegarder les documents dans une base locale. Le but de ce module est le téléchargement de toutes les URL trouvées.

Cette étape nécessite une connexion haut débit pour accélérer le fonctionnement de ce module. Les deux modules 'Recherche et Téléchargement' peuvent fonctionner en simultanés. Chaque URL trouvée lors de la recherche lance un appel au module Téléchargement. Ce dernier peut exécuter plusieurs tâches à la fois, ce qui nous donne un temps de réponse très réduit avec bien entendu une bonne connexion.

Le téléchargement des documents s'effectue uniquement dans le serveur de l'application. Cette étape représente un prétraitement par rapport aux étapes de filtrage des documents, d'extraction des sous fragments et de fusion, enfin d'adaptation aux profils des apprenants.

L'analyse de ces documents est détaillée dans l'étape suivante.

7.5.3. Analyser le contenu des informations trouvées

Analyser le contenu des informations trouvées revient à analyser le contenu des documents téléchargés dans la base locale. Ce qui est visé dans ce deuxième filtrage, c'est le contenu des documents. Google permet de rechercher aussi bien des références de documents que les documents eux-mêmes ou encore le contenu de ces derniers. La recherche est basée sur la requête formulée dans PrOALDoF-Web, ce qui implique une classification non effectuée des documents par rapport aux concepts de l'ontologie de domaine, aussi bien du contenu des documents par rapport aux sous concepts de la même ontologie. Car la recherche effectuée est reliée au sujet général avec ses termes globaux.

Nous voulons par cette étape une association de chaque document à un ou plusieurs concepts de l'ontologie. En PrOALDoF-Web, le sujet de recherche présente un cours par la suite, et chaque concept de l'ontologie présente un chapitre.

En général, si nous recherchons des documents pour un cours donnée, les résultats peuvent englober tout le sujet de recherche, ou bien, une partie du sujet. Pour cela, nous devons faire des statistiques pour accomplir cette association.

Le filtrage dans ce contexte est basé sur une méthode statistique, ce qui permet d'obtenir forcément des répétitions d'informations dans les fragments des documents. La sélection des bons fragments devient une tâche très délicate avec peut être une incertitude d'englober tout le besoin du cours à construire, parce que nous abordons des notions de sémantique d'un texte et qu'une méthode statistique ne le permet pas.

Pour cela, une deuxième association plus profonde est nécessaire, elle vise à associer chaque sous fragment d'un document au sous concept de l'ontologie lui correspondant. Cette étape vient d'analyser le contenu sémantiquement en se basant sur la distance entre les termes des phrases. Les résultats de cette analyse vont être sauvegardés dans la base des nouveaux fragments 'BNF'.

A la fin de cette étape de recherche, la base 'BNF' contient tous les sous fragments correspondants aux concepts de l'ontologie (plus précisément, aux règles sémantiques de chaque concept).

7.6. Représentation des connaissances 'Ontologie de domaine'

Nous utilisons une ontologie de domaine pour présenter les informations des cours sémantiquement. Les ontologies en un rôle très important comme moyen solide de représentation de connaissances.

7.6.1. Représentation est un substitut

Les objets physiques, les événements et les relations qui ne peuvent pas être stockés directement dans un ordinateur, sont représentés par des symboles qui servent de remplaçants. Les symboles et les liens entre eux forment un modèle du système externe. En manipulant les remplaçants internes, un programme computationnel peut simuler le système externe ou raisonner à son sujet.

7.6.2. Représentation est un ensemble d'engagements ontologiques

L'ontologie est l'étude de l'existence. En ingénierie des connaissances, pour un système à base de connaissances, l'ontologie détermine les catégories de concepts qui existent ou qui peuvent exister dans un domaine d'application. Ces catégories représentent les engagements ontologiques des concepteurs ou des ingénieurs de la connaissance.

7.6.3. Représentation est une théorie fragmentaire du raisonnement intelligent

Pour soutenir le raisonnement à propos des choses d'un domaine, une représentation des connaissances peut aussi décrire leurs attitudes et leurs interactions. La description constitue une théorie du domaine d'application. La théorie peut être établie en axiomes explicites, ou peut être compilée en programmes exécutables.

7.7. Description du Domaine d'application

Le squelette constitue la structure du corps humain et protège les organes internes et délicats, après sa naissance, un bébé possède en moyenne 350 os, la plupart de ces os fusionnent quand il grandit (Guéuard, 2002). Un adulte n'a en moyenne que 206 os, à la naissance le crâne a 26 os séparés.

Les os sont connectés par des membranes souples et cartilagineuses appelées 'FENTANELE', après quelque temps la membrane s'ossifie et crée les sutures rigides du crâne adulte (Chung, 1996).

Les os du crâne forment une caisse dure et protectrice qui protège votre cerveau et d'autres organes, les os sont liés à votre colonne vertébrale qui est composée de 32 petits os appelés 'vertèbres', la colonne vertébrale est l'épine dorsale de votre corps, tous les autres os sont directement ou indirectement attachés, chaque os dans le corps a une forme et une grandeur unique selon ses fonctions et son emplacement, 12 paires de côtes sont maintenues à la colonne vertébrale, la cage formée par les côtes protège les organes principaux, les côtes qui sont solides mais souples montent et descendent à la respiration (Guéuard, 2002).

Votre main et les os qu'ils l'entourent constituent un tour de force impressionnant, elle est constituée de 27 os, les articulations complexes de la main, du coude, et de l'épaule sont telles qu'elles permettent un vaste champ de mouvement, le bras est bien adapté aux manipulations (Khelil, 1994).

Le pelvis (ceinture pelvienne) est la structure osseuse située dans la partie inférieure de votre colonne vertébrale, il lie vos jambes au squelette (Guéuard, 2002).

Les os de vos jambes constituent les supports de votre corps, ils sont bien plus grands et solides que les os des bras, l'os de la cuisse appelé 'FEMUR' est l'os le plus long de tout le corps (Guéuard, 2002).

Chaque pied est composé de 26 os, ils permettent une grande flexibilité et soutiennent le poids de tout votre corps, quand vous êtes debout, le poids est distribué sur le long des pieds, la moitié du poids repose sur le 'CALCANEUM' (l'os du talon), et l'autre moitié repose sur les os du tarse, qui composent le pied. Les os sont parfaitement adaptés à leur fonction (Guéuard, 2002).

Voir Figure 11 ci-dessous qui présente un extrait hiérarchique de la partie 'Cou et Tronc' de notre ontologie de domaine créée dans le cadre de notre projet PrOALDoF-Web.

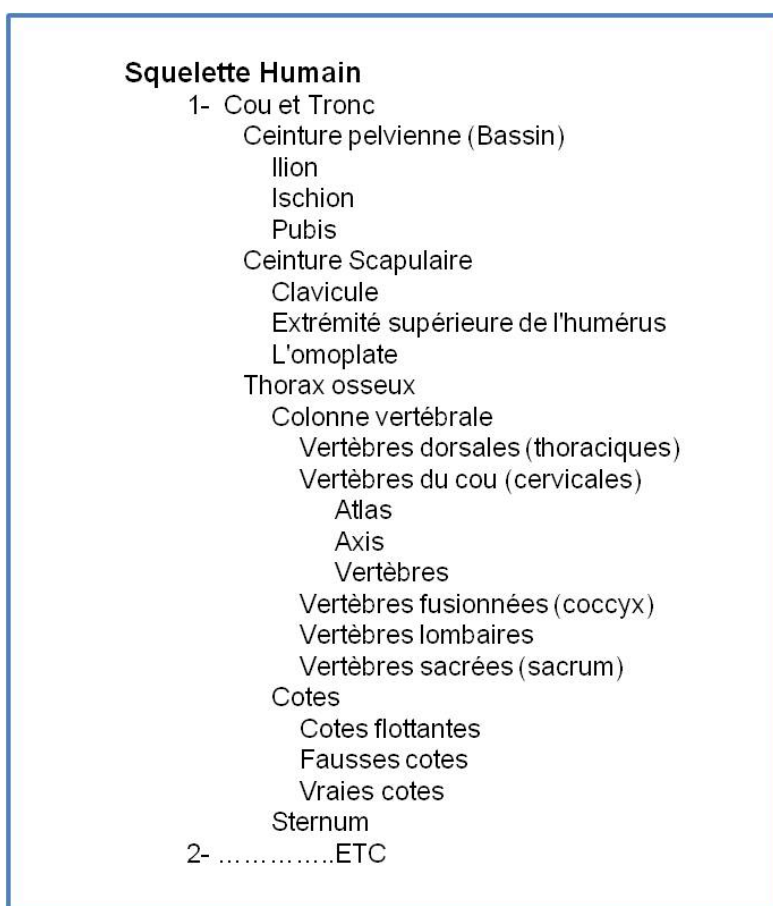


Figure 11 Extrait hiérarchique du 'Squelette Humain'

7.8. La méthodologie ontologique suivie : Uschold et King's

Nous avons suivi la méthodologie de Uschold et Kings, pour construire une ontologie sur l'anatomie humaine et nous avons choisi une sous branche de ce domaine la 'le squelette humain', toute information a été extraite à partir des trois livres (Chung, 1996), (Guéuard, 2002) et (Khelil, 1994). Le choix de la méthodologie était à la condition que cette dernière répond bien à notre objectif final.

La méthodologie comporte 4 étapes :

- (1) : Identification des objectifs et du contexte ;
- (2) : Construction de l'ontologie ;
- (3) : Evaluation de l'ontologie ;
- (4) : Documentation de l'ontologie.

Les quatre étapes sont obligatoires pour créer une ontologie complète pour laquelle répond aux besoins de la formation en ligne demandée.

La Figure 12 ci-dessous présente un extrait de l'ontologie 'Anatomie Humaine' créée sous l'outil protégé2000.

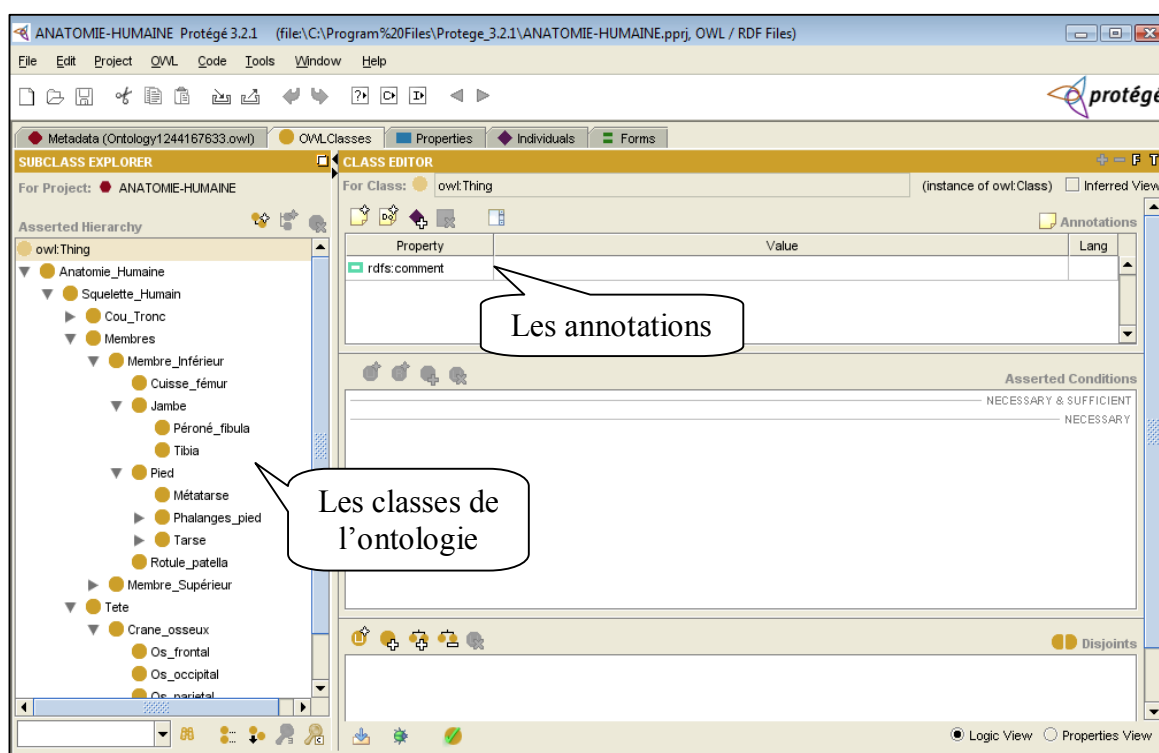


Figure 12 Un extrait de l'ontologie 'Squelette Humain' sous Protégé2000

7.8.1. Identification des objectifs et du contexte

Notre ontologie est construite pour représenter les connaissances du squelette humain sous forme hiérarchique dont le but de faire un lien entre le thème recherché par PrOALDoF-Web et les documents téléchargés, c'est-à-dire, les concepts de l'ontologie vont être les chapitres et les sous titres du cours à réaliser.

Ainsi, les annotations de chaque concept présentent un moyen de comparaison pour les documents téléchargés, afin d'extraire les parties qui ont le plus grand degré de pertinence dans chaque document.

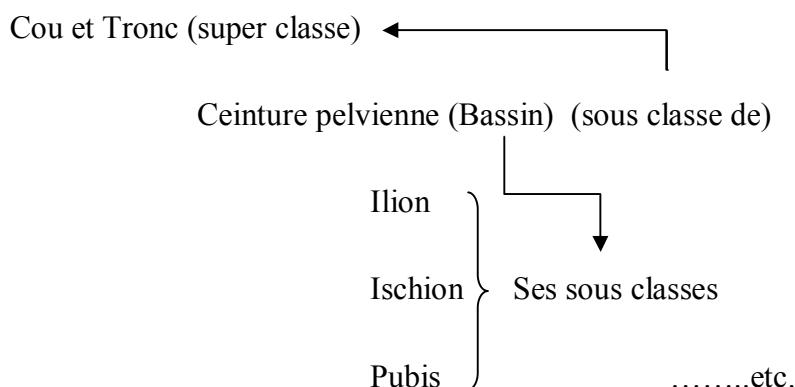
7.8.2. Construction de l'ontologie

Cette étape est divisée en trois activités :

- (1) : Capture de l'ontologie ;
- (2) : Codage de l'ontologie ;
- (3) : Intégration d'ontologies existantes.

7.8.2.1. Première activité : 'capture de l'ontologie'

Dans cette activité nous avons identifié tous les concepts (classes) et leurs slots (attributs) et aussi comment les instanciés. Par exemple, pour la classe père squelette, nous avons 3 sous classes : Tête, Membres, Cou et tronc. Chaque partie (classe) a aussi des sous classes ; par exemple :



Ainsi de suite jusqu'à l'identification de toute la hiérarchie du squelette ; bien sur tout ça sur papier. La Figure 11 en page 117 montre un extrait de cette hiérarchie, et pour mieux comprendre la Figure 13 ci-dessous montre un extrait graphique en utilisant l'outil Protégé2000.

Pour les slots, nous avons identifié neuf (09) au total par exemple : Type d'os, Nombre d'os ...etc. Nous avons associé chaque slot aux concepts correspondants avec les valeurs. Par exemple, le nombre d'os est de la valeur 206 pour la classe : squelette. Donc, l'approche utilisée est l'**approche descendante**, c'est-à-dire à partir de concepts abstraits que l'on spécialise en concepts plus spécifiques.

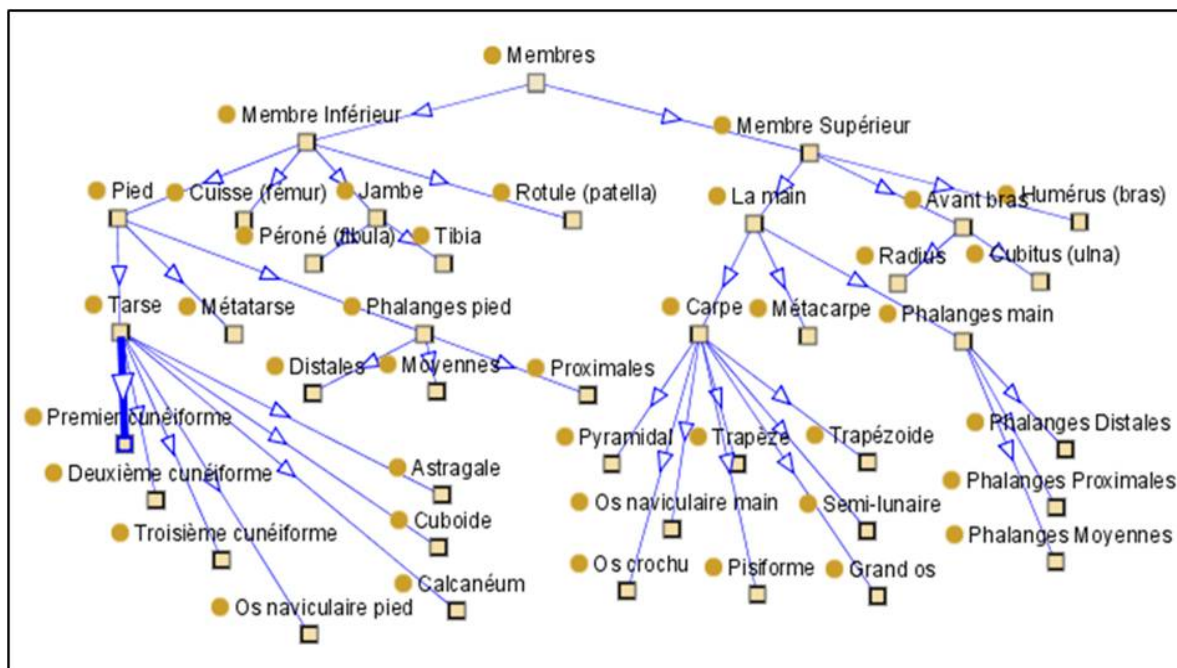


Figure 13 Un extrait graphique de l'ontologie de domaine créée 'Squelette Humain'

7.8.2.2. Deuxième activité : 'codage de l'ontologie'

Le codage a commencé après la terminaison de toute la hiérarchie de l'ontologie ; l'outil utilisé est : Protégé 2000. Ce codage permet la création des concepts, des slots et des instances.

7.8.2.3. Troisième activité : 'intégration d'ontologies existantes'

C'est une nouvelle ontologie créée spécialement pour notre besoin. Donc, nous ne sommes pas besoin d'une ou de plusieurs ontologies existantes.

7.8.3. Evaluation de l'ontologie

Après le codage de toute l'ontologie, un médecin généraliste 'DOCTEUR : KOUCHOUK.F' a évalué notre ontologie, il a trouvé que la hiérarchie est bonne et elle peut répondre à notre besoin qui est la construction d'un cours 'chapitres et sous chapitres'.

Après avoir évalué la hiérarchie, il a commencé à vérifier les slots et surtout la notion d'héritage de ces derniers puisque ces slots présentent des mots-clés par la suite dans le travail. Pour les instances, ce sont des valeurs que nous donnons pour chaque concept, et plus précisément pour chaque slot d'un concept, l'évaluation du même docteur était valide.

7.8.4. Documentation de l'ontologie

Cette étape consiste à documenter l'ontologie, nous la faisons après la terminaison de la création de toutes les classes et les slots, mais pour notre objectif du travail les documentations représentent des annotations. Donc, nous avons choisi de créer des mots-clés pour chaque concept au lieu d'un texte.

7.9. Intégration de l'ontologie de domaine

Cette intégration suit un modèle proposé dans PrOALDoF-Web en utilisant l'API JENA. Il faut d'abord tirer la hiérarchie du cours à partir de l'ontologie de domaine avant de commencer le traitement. Puis, transformer l'ontologie à un fichier extensible à tout changement et mise à jour lors du filtrage des documents, extraction et sélection des ressources pertinentes pour chaque partie du cours.

7.9.1. L'API JENA

Nous avons utilisé l'API JENA qui ne fonctionne que dans le langage JAVA. L'API JENA donne une possibilité d'intégration de l'ontologie de domaine dans le système (Voir un extrait graphique de l'ontologie réalisée au sein du projet PrOALDoF-Web dans la Figure 13 en page 120).

JENA est un Framework Java pour la construction sémantique des applications Web. Il fournit un environnement de programmation pour RDF, RDFS et OWL, SPARQL (langage de requêtes pour RDF) et comprend une règle basée sur un moteur d'inférence.

JENA est open source. Elle est développée à partir des travaux réalisés dans le laboratoire centrale de recherche de HP 'Hewlett-Packard', dans le domaine 'Recherche sur le Web sémantique' ou 'Semantic Web Research'.

Avec cette API, nous pouvons lire et écrire notre ontologie de domaine 'Squelette Humain' de type OWL (OWL). La hiérarchie du cours est obtenue ; en utilisant cette API ; à partir l'ontologie.

7.9.2. La hiérarchie du cours

Nous extrayons la hiérarchie du cours à partir de l'ontologie de domaine (Voir Tableau 3 en page 122).

Ontologie de Domaine			Le cours		
Concept 1	Concept 2	Au Concept L	Chapitre 1	Chapitre 2	Au Chapitre A
Sous Concept 1	Sous Concept 2	Au Sous Concept M	Sous Titre 1	Sous Titre 2	Au Sous Titre B
L : Le dernier Concept de l'ontologie ; M : Le dernier Sous Concept du Concept père correspondant.			A : Le dernier Chapitre du Cours ; B : Le dernier Sous Titre du Chapitre père correspondant.		

Tableau 3 La Hiérarchie du cours à partir de l'ontologie de domaine

L'ontologie se transforme en fichier Excel présentant une base de sauvegarde de toutes les parties sélectionnées à partir des documents à filtrer.

7.9.3. Transformation de l'ontologie en fichier Excel

Avant de commencer le filtrage des documents téléchargés, il faut d'abord penser à: *comment implémenter l'ontologie dans le système ?*

Nous avons l'outil nécessaire, il nous manque la manière d'implémentation. Alors, une simple idée était l'extraction des concepts, slots et instances. Nous utilisons l'API JENA pour achever l'extraction. Mais, *est-ce que nous pouvons garder la hiérarchie de l'ontologie (du cours) après l'extraction des concepts ?* (Voir le principe d'une hiérarchie ontologique sous l'outil Protégé 2000 dans la Figure 14 ci-dessous)

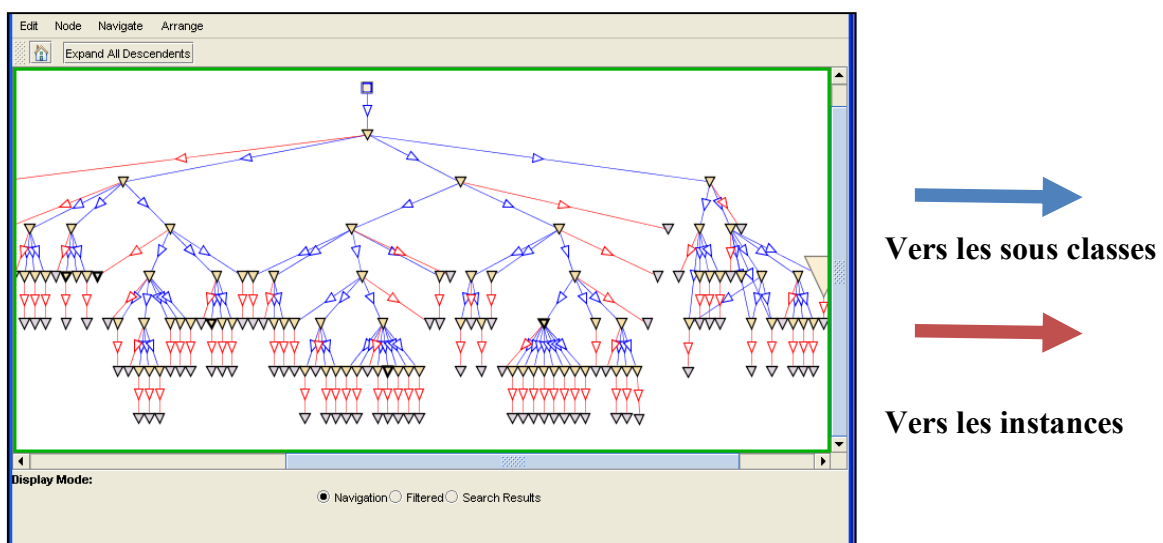


Figure 14 Extrait d'une ontologie sous Protégé2000

Avec la classe de JENA responsable en extraction de la hiérarchie d'une ontologie, nous ne pouvons pas distinguer entre les concepts pères et leurs fils car l'ontologie est un graphe. Par contre, nous voulons garder la hiérarchie de l'ontologie dans notre cours à créer (Voir Figure 15 ci-dessous).

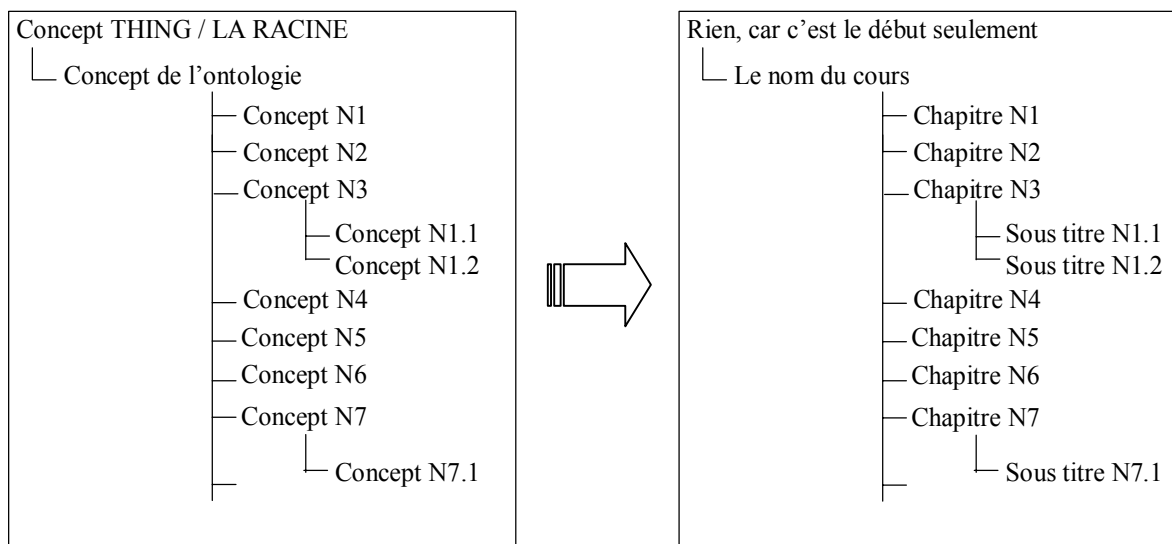


Figure 15 La transformation de l'ontologie sous le format Excel

Cette transformation doit être assurée pour que le cours soit complet d'une part. D'une autre part, les annotations de chaque concept doivent être aussi insérées dans le modèle proposé, car, elles présentent par la suite, un moyen de comparaison dans le filtrage des informations du Web.

De plus, nous voulons réaliser un historique de toutes les parties pertinentes extraites pour chaque chapitre et chaque sous titre du cours. Cet historique augmente l'efficacité du système en proposant un moyen de contrôle très puissant avec de nouvelles propositions aux concepteurs des cours.

Pour ces deux raisons, notre modèle Excel devient une nécessité dans PrOALDoF-Web. Nous avons utilisé l'API Excel Java pour sauvegarder les résultats dans un fichier Excel.

Cette API est un kit de développement logiciel, Open Source, permettant de lire, d'écrire, et de modifier dynamiquement les feuilles de calcul Excel. Maintenant les développeurs Java peuvent lire des feuilles de calcul Excel, de les modifier avec une API simple et commode, et écrire les changements à n'importe quel flux de sortie (par exemple : disque, HTTP, base de données, ou toute autre prise de courant).

Tout système d'exploitation qui peut exécuter une machine virtuelle Java (à savoir, pas seulement Windows) peut à la fois développer et atteindre les feuilles de calcul Excel. Parce qu'il est de Java, l'API peut être appelée dans une Servlet, donnant ainsi accès à des feuilles de calcul Excel, grâce à internet et intranet 'Applications Web'.

Dans PrOALDoF-Web, notre besoin de ce modèle, est de sauvegarder tous les résultats obtenus ; en filtrant les ressources Web téléchargés ; localement.

7.10. Filtrage des informations trouvées

Le filtrage des informations sauvegardées est basé essentiellement sur deux méthodes de calculs. La première est appelée : Degré de Pertinence 'DP', et la seconde appelée : Distance Basée sur les Règles Sémantiques 'DBRS'. Les deux méthodes fonctionnent en parallèle pour filtrer tous les documents téléchargés à partir du Web. Le filtrage de la base de documents, permettre de stocker chaque partie pertinente extraite dans un fichier Excel qui présente une base de stockage des résultats du filtrage.

L'extraction d'informations cherche à analyser de manière précise le contenu d'un document, contrairement à la recherche de document qui étudie sa thématique générale. Il s'agit donc, d'une tâche qui ne peut être accomplie qu'après des sélections préalables dans les documents, et qui est considérée comme plus ardue que la catégorisation de documents. A ce stat, une suite de calculs qui doit s'effectuer pour pourvoir sélectionner et extraire les bonnes parties de chaque document. *Alors, comment faire l'extraction d'informations ?*

7.10.1. Extraction d'informations

Avant d'extraire les informations pertinentes, il faut d'abord, effectuer une analyse morpho-lexicale sur les documents en identifiant les phrases et les mots du texte de chaque document.

En premier lieu, nous découpons chaque document en parties ou paragraphes. Ce découpage consiste à numéroter les parties du document comme suit :

<P1> la partie 1 du document <P1/>

<P2> la partie 2 du document <P2/>

.....

<PN> la partie N du document <PN/>

Où N est le nombre maximal des parties du document.

Après, nous faisons une segmentation pour chaque partie, dont le but est de découper le texte en phrases puis en mots distincts (Voir la Figure 16 ci-dessous qui montre un exemple de texte segmenté d'un extrait d'un fichier PDF).

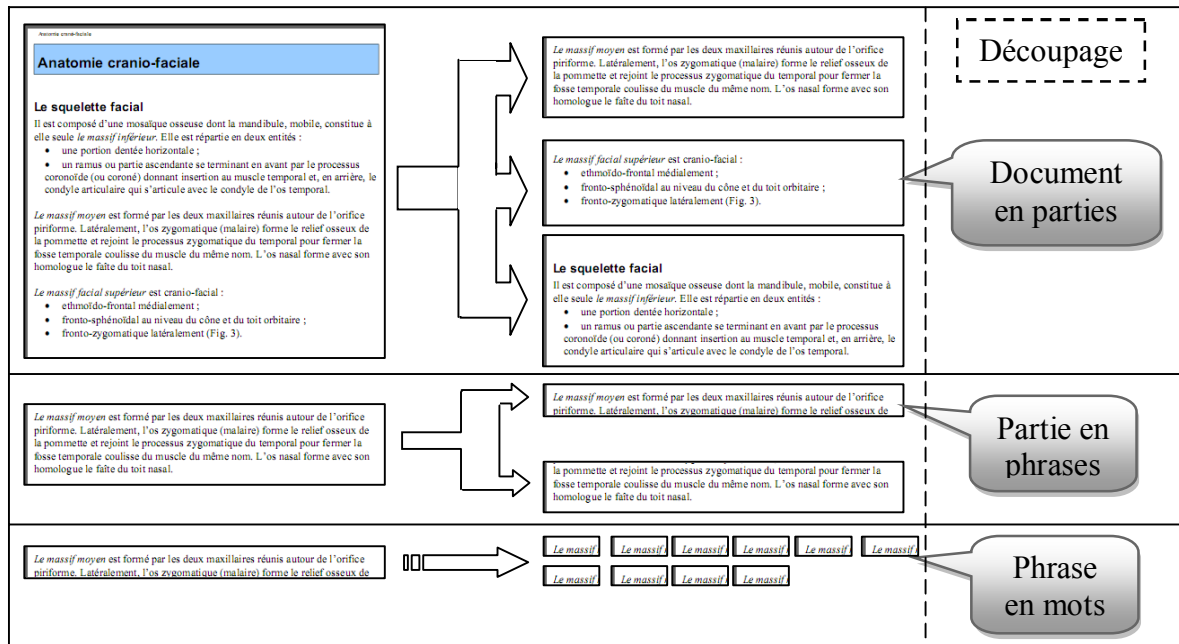


Figure 16 Extrait d'un document PDF segmenté

Cette segmentation en phrases pour la partie une 'P1', puis en mots pour la phrases 1 'PH1' est comme suit :

<S1> la phrase 1 de P1 <S1/>

<M1> le mot 1 de S1 <M1/>

<M2> le mot 2 de S1 <M2/>

.....

<MN> le mot N de S1 <MN/>

Où N est le nombre maximal des mots de S1.

<S2> la phrase 2 de P1 <S2/>

.....

<SN> la phrase N de P1 <SN/>

Où N est le nombre maximal des phrases de P1.

Nous numérotons toute information (Partie, Phrase et Mot) de chaque document, pour faciliter la programmation dans l'étape de filtrage, surtout, les calculs des deux méthodes 'DP' et 'DBRS'.

Après la segmentation, nous éliminons les mots inutiles (le, la, de, du, ...etc.), pour que le résultat du degré de pertinence ne sera pas influencé par le nombre important de ces mots inutiles. La Figure 16 en page 125 résume le processus du traitement d'un document dans PrOALDoF-Web afin d'obtenir des mots séparés. Nous ne faisons pas les autres étapes de l'analyse (la lemmatisation et l'étiquetage), *pourquoi ?*

Notre méthode de filtrage est basée sur la statistique en calculant un degré de pertinence ou encore une distance entre les termes des phrases. Donc, notre besoin est juste la localisation de tous les termes de chaque partie de chaque document téléchargés. De plus, nous ne voulons pas savoir les formes canoniques des mots (s'ils sont des verbes, noms ...etc.), car les concepts et leurs mots-clés sont en réalité des mots techniques qui sont définies par des experts dans le domaine.

Dans PrOALDoF-Web l'extraction d'informations devient une extraction des termes. Ces derniers vont être comparés avec les concepts et les mots-clés extraits à partir de l'ontologie de domaine, afin d'extraire les parties pertinentes de chaque partie du document.

7.10.2. Degré de Pertinence 'DP'

$$DP = \left(F_c + \sum_{mc=0}^n (F_{mc} * P_{mc}) \right) / N \quad (1)$$

Formule 1 Degré de Pertinence 'DP'

Où :

c : Un concept de l'ontologie de domaine ;

mc : Un mot-clé d'un concept ;

F_c : La fréquence d'un concept dans un seul fragment d'un document Web ;

F_{mc} : La fréquence d'un mot-clé dans un seul fragment d'un document Web ;

n : Le nombre maximum des mots clés ;

P_{mc} : Le poids d'un mot-clé ;

N : Le nombre maximum des mots dans un seul fragment d'un document Web ;

$\sum_{mc=0}^n (F_{mc} * P_{mc})$: La somme des fréquences de tous les mots-clés ($mc=0 \dots n$) d'un seul concept dans un seul fragment d'un document Web, chaque fréquence est multipliée par le poids du mot-clé correspondant.

Le degré de pertinence 'DP' est une méthode statistique (1), sur la base de la fréquence du concept de l'ontologie (qui présente un composant de la formation) dans un fragment d'une ressource Web d'une part, et l'existence des mots-clés et de leurs fréquences dans le même fragment d'une autre part.

La fréquence du mot dans un fragment est : combien de fois le mot peut être répété dans ce fragment. La Formule 1 en page 126 permet de calculer le 'DP' dans un fragment d'un document Web.

Pour dire qu'un fragment est pertinent pour une composante d'un cours, nous calculons le degré de pertinence de ce fragment par rapport à toutes les composantes de ce cours. Le fragment qui a le plus grand degré sera sauvegardé dans la base 'BNF', puis nous répétons le même processus pour tous les fragments des ressources Web (documents) trouvés.

Exemple : le calcul du degré de pertinence pour deux fragments par rapport à deux composantes du cours

Un fragment du document = mot1 mot2 mot1 mo3 mo4 mo1 mo2 mot4 mot5 mot1

La composante 1 = mot1

Le composante 2 = mot2

Les mots-clés = mot3, mot5

Les mots-clés = mot4, mot5

$$DP_{\text{mot1}} = 0.6$$

$$DP_{\text{mot2}} = 0.5$$

D'après les deux résultats, ce fragment Web doit être associé à la composante 1. C'est-à-dire, ce fragment est un fragment pertinent pour la composante du cours mot1. Nous sauvegardons ce fragment avec son 'DP' dans le fichier Excel du modèle proposé. Nous détaillons ce modèle dans le chapitre suivant.

Nous répétons ce traitement pour tous les fragments d'un document, afin de le compléter. Maintenant, si nous trouvons deux ou plusieurs fragments pertinents pour une même composante du cours, *qu'est que nous faisons ?*

Selon la méthode 'DP', nous aurons deux degrés pour les fragments jugés pertinents pour la même composante du cours. Pour cela, nous devons calculer la moyenne des degrés de pertinences des fragments en situation (2).

La Formule 2 en page 128 montre comment calculer le degré de pertinence pour plusieurs fragments jugés pertinents dans le même document pour une seule composante du cours.

$$DPM = \left(\sum_{f=0}^{nf} (DP_f) \right) / nf \quad (2)$$

Formule 2 Degré de Pertinence Moyen ‘DPM’

Où :

f : Un fragment d’un document Web ;

nf : Le nombre maximum des fragments jugés pertinents pour la même composante du cours ;

DP_f : Le Degré de Pertinence pour le fragment f

$\sum_{f=0}^{nf} (DP_f)$: La somme de tous les fragments jugés pertinents pour la même composante du cours ;

Si nous terminons le traitement du calcul du degré de pertinence d’un document, nous devons répéter le même traitement pour tous les documents téléchargés dans la base locale.

Enfin, le but de cette étape ‘Filtrage’, est l’extraction des fragments pertinents qui se trouvent dans les documents téléchargés à partir du Web. Ces fragments vont être classés en ordre selon le degré de pertinence dans un fichier Excel ‘modèle proposé’ pour chaque composante du cours.

7.10.3. Distance Basée sur les Règles Sémantiques ‘DBRS’

Il s'agit d'une méthode sémantique composée de trois formules : (3), (4) et (5) ‘DBRS_M’, ‘DBRS_D’ et ‘DBRS_{PA}’, basée sur la distance entre les mots dans les sous fragments (phrases). Tout d'abord, nous devons extraire les mots d'un sous fragment en appliquant la méthode d'extraction d'informations précédemment détaillée. Puis, nous calculons la distance seulement entre les mots par rapport aux termes qui sont définis dans les règles sémantiques ‘RS’ pour chaque composante du cours. ‘DBSR’ présente une projection des ‘RS’ sur les sous fragments des ressources Web trouvées, pour extraire les sous fragments les plus pertinents d’une partie du cours. Lorsque nous avons des mots successifs dans les sous fragments des documents Web, successifs par rapport aux termes des règles sémantiques. Dans ce cas, quelques distances seront nulles ‘0’ et la distance ‘DBRS_D’ rapproche du 0. En conséquence, la précision d’association ‘DBRS_{PA}’ sera la plus haute par rapport aux autres précisions. Dans ce cas, nous devons associer le sous fragment à la règle sémantique en question.

Dans le cas où le système ne trouve aucun sous fragment pour une composante du cours, nous devons utiliser un dictionnaire des termes synonymes aux termes des ressources Web, pour donner plus de possibilités d'obtenir les sous fragments demandés qui contiennent les différentes informations définies dans les 'RS'.

La Formule 3 ci-dessous montre comment calculer la moyenne 'DBRS_M' ; avant le calcul de la distance 'DBRS_D' ; pour un sous fragment d'un document Web.

$$DBRS_M = \left(\sum_{t=0}^n (Si E_t \text{ dans } SF \text{ Alors } 1 \text{ Sinon } E_{ts}^{tsn} \text{ dans } SF) \right) / N \quad (3)$$

Formule 3 Distance Basée sur les Règles Sémantiques (Moyenne) 'DBRS_M'

Où :

t : Un terme dans une règle sémantique ;

n : Le nombre maximum des termes dans une règle sémantique ;

SF : Un sous fragment ;

E_t : L'existence du terme t dans le sous fragment SF ;

ts : Un terme synonyme du terme t ;

tsn : Le nombre maximum des termes synonymes du terme t ;

E_{ts}^{tsn} : L'existence du terme synonyme ts dans le sous fragment SF ;

$\sum_{t=0}^n (Si E_t \text{ dans } SF \text{ Alors } 1 \text{ Sinon } E_{ts}^{tsn} \text{ dans } SF)$: La somme de tous les termes d'une règle sémantique dans un sous fragment ;

N : Le nombre maximum des mots dans un sous fragment ;

Le but de la première partie de la méthode 'DBRS_M' est de marquer tous les termes existants dans un sous fragment par rapport aux termes de la règle sémantique, dont le but est de trouver la moyenne de l'existence des termes d'une règle sémantique dans un sous fragment.

Si cette moyenne rapproche du 1, alors, il est fort possible que le sous fragment relève de cette règle sémantique, car les distances entre les termes dans ce cas sont toutes nulles.

Dans le cas contraire, nous devons calculer la distance entre les termes existants dans un sous fragment, puis nous comparons toutes les distances par rapport aux règles sémantiques. La distance la plus petite est celle qui présente la meilleure association du sous fragment à cette règle sémantique qui a cette distance.

Dans la formule (3), le but est de marquer tout terme trouvé. Autrement dit, le but est de trouver l'indice de chaque terme dans le sous fragment en cours de l'analyse, c'est-à-dire l'ordre du terme par rapport aux mots du sous fragment. Si les termes sont tous trouvés, nous pouvons appeler la formule (4) pour calculer la distance entre les termes dans le sous fragment.

Nous jugeons qu'un sous fragment est pertinent pour une règle sémantique particulière, si la distance 'DBRS' est la plus petite pour cette règle par rapport aux règles sémantiques existantes pour cette composante du cours.

La Formule 4 ci-dessous montre comment calculer la distance 'DBRS' pour un sous fragment d'un document Web.

$$DBRS_D = \sum_{i=1}^{n-1} D_{(m+1)-m} \quad (4)$$

Formule 4 Distance Basée sur les Règles Sémantiques 'DBRS_D'

Où :

i : Un indice d'un terme dans un sous fragment ;

n : Le nombre maximum des indices trouvés dans un sous fragment ;

$D_{(m+1)-m}$: La distance entre deux indices, c'est-à-dire, le nombre de mots entre les deux termes ayant l'indice m et l'indice suivant $m + 1$, un indice d'un terme composé de plusieurs termes égale à la moyenne de ces derniers ;

$\sum_{i=1}^{n-1} D_{(m+1)-m}$: La somme de toutes les distances entre les indices trouvés.

7.10.4. Précision d'une association 'PA'

C'est une mesure dont le but est d'assurer l'association d'un sous fragment à une seule règle sémantique. La Formule 5 en page 131 montre comment calculer la précision d'une association 'PA' pour un sous fragment d'un document Web par rapport à une règle sémantique.

Nous devons répéter le même traitement ; calcul des moyennes et distances pour tous les sous-fragments des ressources Web, de les associer aux règles sémantiques qui représentent toutes les composante du cours, et enregistrer les résultats dans la base 'BNF'. Après cela, nous devons réaliser le processus de fusion avec des sous fragments extraits à partir des ressources Web (Voir Chapitre III Partie II en page 132).

$$DBRS_{PA} = DBRS_M * \left(\frac{NB_{te}}{DBRS_D} \right) \quad (5)$$

Formule 5 Précision d'une association 'PA' 'DBRS_{PA}'

Où :

DBRS_M: La moyenne des termes d'une règle sémantique dans un sous fragment ;

NB_{te} : Nombre des termes existants dans un sous fragments ;

DBRS_D : Distance Basée sur les Règles Sémantiques entre les termes existants dans un sous fragment ;

$\frac{NB_{te}}{DBRS_D}$: La division du nombre des termes existants dans un sous fragments par la distance entre ces termes ;

$DBRS_M * \left(\frac{NB_{te}}{DBRS_D} \right)$: La multiplication de la moyenne des termes existants par rapport aux mots du sous fragment 'DBRS_M' avec le résultat de la division du nombre des termes existants dans un sous fragment 'NB_{te}' par la distance entre ces termes dans le même sous fragment 'DBRS_D'.

7.11. Conclusion

Nous avons présenté à travers ce chapitre notre méthode de recherche dans le Web pour trouver les informations les plus pertinentes via le moteur de recherche Google. Puis, nous avons vu comment filtrer chaque ressource Web trouvée dont le but est l'extraction des informations adéquates au besoin de la formation.

Nous avons présenté nos formules de calculs pour trouver les sous fragments les plus pertinents. Pour chaque sous fragment, nous devons calculer le degré de pertinence 'DP' et la distance basée sur les règles sémantiques 'DBRS'. Cette dernière est composée de trois méthodes 'DBRS_M' (Moyenne), 'DBRS_D' (Distance) et 'DBRS_{PA}' (Précision de l'Association).

Dans le chapitre suivant, nous détaillons comment fusionner les ressources Web trouvées. Puis, comment les adaptées aux profils des apprenants dans le cadre de notre projet PrOALDoF-Web.

8. La Fusion et l'Adaptation dans PrOALDoF-Web

8.1. Introduction

Nous voulons par ce chapitre, la présentation de nos deux approches essentielles dans notre projet PrOALDoF-Web.

La première approche est la fusion des ressources Web jugées pertinentes. L'approche utilise un modèle proposé et des méthodes statistiques et sémantiques pour sauvegarder tous les résultats trouvés. Nous détaillons le principe de la fusion avec ses différentes étapes afin d'obtenir le but fixé.

La deuxième approche est l'adaptation des ressources pertinentes aux profils des apprenants en se basant sur une ontologie pédagogique.

8.2. La Fusion dans PrOALDoF-Web

C'est un processus automatique basé sur des méthodes statistiques et sémantiques. Le principe de la fusion dans PrOALDoF-W, est de combiner plusieurs informations ; jugées pertinentes à un domaine particulier ; en une seule entité. Les informations sont extraites à partir nombreuses sources d'informations Web qui peuvent être hétérogènes et qui ont des formats différents. Chaque information est extraite à partir des textes semi-structuré ou non structuré.

Le but de la fusion est d'améliorer la précision et la qualité des informations combinées, afin d'augmenter les performances du système PrOALDoF-Web et de fournir des informations pertinentes et adéquates aux besoins de la formation. Ainsi, elle tente à pallier les faiblesses de certaines sources d'informations avec les points forts d'autres.

8.3. Les étapes de la Fusion

Nous avons parlé dans le chapitre précédent 'La Recherche et la représentation des connaissances dans PrOALDoF-Web' de la recherche et du filtrage des informations qui se trouvent sur le Web. Ces deux étapes sont basées sur des méthodes statistiques et sémantiques dont le but est d'extraire les informations les plus pertinentes. Nous avons un modèle Excel extensible à tout changement et à toute mise à jour lors l'extraction des informations.

8.3.1. Créer le modèle Excel

Nous devons suivre le principe de la hiérarchie du cours détaillé précédemment pour créer notre modèle Excel (Voir Tableau 4 ci-dessous). La proposition est de réaliser un modèle sous format de table où chaque ligne présente une règle sémantique d'un concept de l'ontologie, chaque concept est une composante du cours.

			DP	DBRS				DP	DBRS				DP	DBRS
Partie 1	RS1	SF1	0	0	SF2	0	0	...	SFN	0	0			
FSF1	RS2	SF1	0	0	SF2	0	0	...	SFN	0	0			
	0	0	...	0	0	0	0			
	RSN	...	0	0	...	0	0	0	0			
Partie 2	RS1	SF1	0	0	SF2	0	0	...	SFN	0	0			
FSF2	RS2	SF1	0	0	SF2	0	0	...	SFN	0	0			
	0	0	...	0	0	0	0			
	RSN	...	0	0	...	0	0	0	0			
...	0	0	...	0	0	0	0			
...	0	0	...	0	0	0	0			
Partie	RS1	SF1	0	0	SF2	0	0	...	SFN	0	0			
FSFN	RS2	SF1	0	0	SF2	0	0	...	SFN	0	0			
	0	0	...	0	0	0	0			
	RSN	...	0	0	...	0	0	0	0			

Tableau 4 Une partie du modèle Excel pour créer la base 'BNF'

La méthode sémantique 'DBRS' est composée de trois formules : (3), (4) et (5) 'DBRS_M' (Moyenne), 'DBRS_D' (Distance) et 'DBRS_{PA}' (Précision de l'association). Donc, dans ce modèle nous avons trois colonnes pour chaque 'DBRS' (Voir Tableau 5 ci-dessous).

			DP	DBRS		
				DBRS _M	DBRS _D	DBRS _{PA}
Partie 1	RS1	SF1	0	0	0	0
FSF1	RS2	SF1	0	0	0	0
	0	0	0	0
	RSN	...	0	0	0	0

Tableau 5 La sauvegarde de la méthode 'DBRS' dans le modèle Excel

En plus, nous avons défini pour chaque composante du cours un ensemble de règles sémantiques 'RS'. Ces dernières sont appelées pour le calcul des trois méthodes 'DBRS_M', 'DBRS_D' et 'DBRS_{PA}' pour chaque sous fragment trouvé dans une ou plusieurs parties d'une ou de nombreuses ressources Web.

Chaque partie du cours est définie et organisée verticalement dans le modèle Excel. Les sous fragments extraits à partir des ressources Web qui sont associés aux règles sémantiques sont sauvegardés dans le modèle Excel et organisés horizontalement (Voir Tableau 4 en page 133).

Nous expliquons les termes dans le Tableau 4 en page 133 comme suit :

- Partie 1 : est une composante du cours ;
- RS1, RS2, ... RSN est l'ensemble des règles sémantiques pour la Partie 1 ;
- SF1, SF2, ... SFN est l'ensemble des sous fragments les plus pertinents pour une règle sémantique de la Partie 1 ;
- DP : Degré de Pertinence d'un sous fragment ;
- DBRS : la Distance Basée sur les Règles Sémantiques d'un sous fragment. La méthode contient les trois formules : DBRS_M' (Moyenne), 'DBRS_D' (Distance) et 'DBRS_{PA}' (Précision de l'association) ;
- FSF1 : Fusion des Sous Fragments. C'est la première partie fusionnée par les premiers sous fragments qui sont ordonnés dans la table pour chaque ligne d'une règle sémantique. FSF : Fusion des Sous Fragments.

Nous présentons un extrait du fichier Excel en appliquant la méthode sur notre ontologie. L'extrait est sur la partie cou et tronc du squelette humain. Cette partie est composée de trois parties (ceinture pelvienne (bassin), ceinture scapulaire et thorax osseux). Dans cet extrait nous focalisons uniquement sur la troisième sous partie de la partie cou et tronc 'thorax osseux' du squelette humain (Voir Tableau 6 en page 135).

Nous utilisons l'API Excel pour la création du modèle Excel. Cette API permet tout changement et toute mise à jour qui peuvent être effectués sur chaque ligne du fichier Excel après l'étape de la fusion. Chaque ligne représente l'ensemble des sous fragments extraits à partir des ressources Web téléchargées localement. Les sous fragments sont ordonnés dans le fichier pour que le choix par la suite pointe directement sur la première ligne.

Squelette Humain						DP	DBRS		DP	DBRS		
Cou et					RS1	SF1	0	0	...	SFN	0	0
FSF					RSN	...	0	0	0	0
	Thorax				RS1	SF1	0	0	...	SFN	0	0
	FSF				RSN	...	0	0	0	0
		Colonne			RS1	SF1	0	0	...	SFN	0	0
		FSF			RSN	...	0	0	0	0
			Vertèbres		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
			Vertèbres du		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
				Atlas	RS1	SF1	0	0	...	SFN	0	0
				FSF	RSN	...	0	0	0	0
				Axis	RS1	SF1	0	0	...	SFN	0	0
				FSF	RSN	...	0	0	0	0
				Vertèbres	RS1	SF1	0	0	...	SFN	0	0
				FSF	RSN	...	0	0	0	0
			Vertèbres		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
			Vertèbres		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
			Vertèbres		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
		Cotes			RS1	SF1	0	0	...	SFN	0	0
		FSF			RSN	...	0	0	0	0
			Cotes		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
			Fausses cotes		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
			Vraies cotes		RS1	SF1	0	0	...	SFN	0	0
			FSF		RSN	...	0	0	0	0
		Sternum		...	RS1	SF1	0	0	...	SFN	0	0
		FSF			RSN	...	0	0	0	0
...	0	0	0	0
...	0	0	0	0
...	0	0	0	0

Tableau 6 Un extrait du modèle Excel d'une portion de la partie Cou et Tronc

Ce modèle Excel répond aux besoins suivants :

- La hiérarchie du cours est la même que celle de l'ontologie ;
- Les annotations et les mots-clés de chaque concept de l'ontologie sont extraits pour jouer deux rôles : le premier, elles servent à générer automatiquement les requêtes de recherche sur le Web. Le second, elles présentent des mots-clés de comparaison dans les méthodes de calculs : 'DP' pour trouver les fragments les plus pertinents pour une composante du cours, 'DBRS' pour trouver les sous fragments les plus pertinents pour chaque sous composante du cours ;
- Enfin, un historique est assuré par notre modèle Excel pour sauvegarder tous les sous fragments extraits à partir des ressources Web pour créer notre base 'BNF'.

Le modèle Excel présente un domaine d'apprentissage, il doit contenir après les étapes d'extraction et de fusion tous les sous fragments et tous les nouveaux fragments via la méthode de fusion, dont le but l'obtention de notre base 'BNF'.

Tout cours construit doit contenir la bibliographie, pour cela nous ajoutons ; à la fin de toute partie extraite à partir des documents, les noms des documents correspondants avec ces URL. Comme ça, par la suite, nous savions la provenance de chaque composante du cours.

8.3.2. Le processus de la Fusion

Avant de commencé la fusion, il y a les étapes de recherche, filtrage et extraction des informations les plus pertinentes qui doivent être effectué. Premièrement, les ressources Web trouvées sont divisées en sections ou paragraphes. Chaque paragraphe est segmenté en phrases 'sous fragments' et en mots séparés en éliminant les mots inutiles (le, de, un, une, ou, dans, ...etc.), afin que ceux-ci n'affectent pas sur les résultats des deux méthodes 'DP' et 'DBRS'.

Pour créer des pièces 'sous fragments' d'un cours, nous devons extraire d'abord les fragments pertinents pour un concept de l'ontologie de domaine 'composante du cours' à partir des ressources Web trouvées. Pour cette raison, nous devons calculer les degrés de pertinence 'DP' dans un premier temps, pour associer un fragment Web à une seule composante du cours. Dans un second temps, nous devons calculer les trois formules de la méthode 'DBSR' seulement sur les fragments les plus pertinents pour extraire des sous fragments qui vont être associés aux règles sémantiques de chaque composante du cours.

Nous pouvons extraire des sous fragments à partir des fragments multiples des ressources Web. Les sous fragments d'une ressource Web peuvent être associés à une seule composante du cours ou bien, ils vont être dispatchés aux plusieurs composantes du cours.

Nous pouvons illustrer le processus de la Fusion en utilisant cet exemple (Voir le principe du processus de Fusion dans la Figure 17 ci-dessous) : La Fusion de trois fragments les plus pertinents dans trois documents différents pour la même partie du cours.

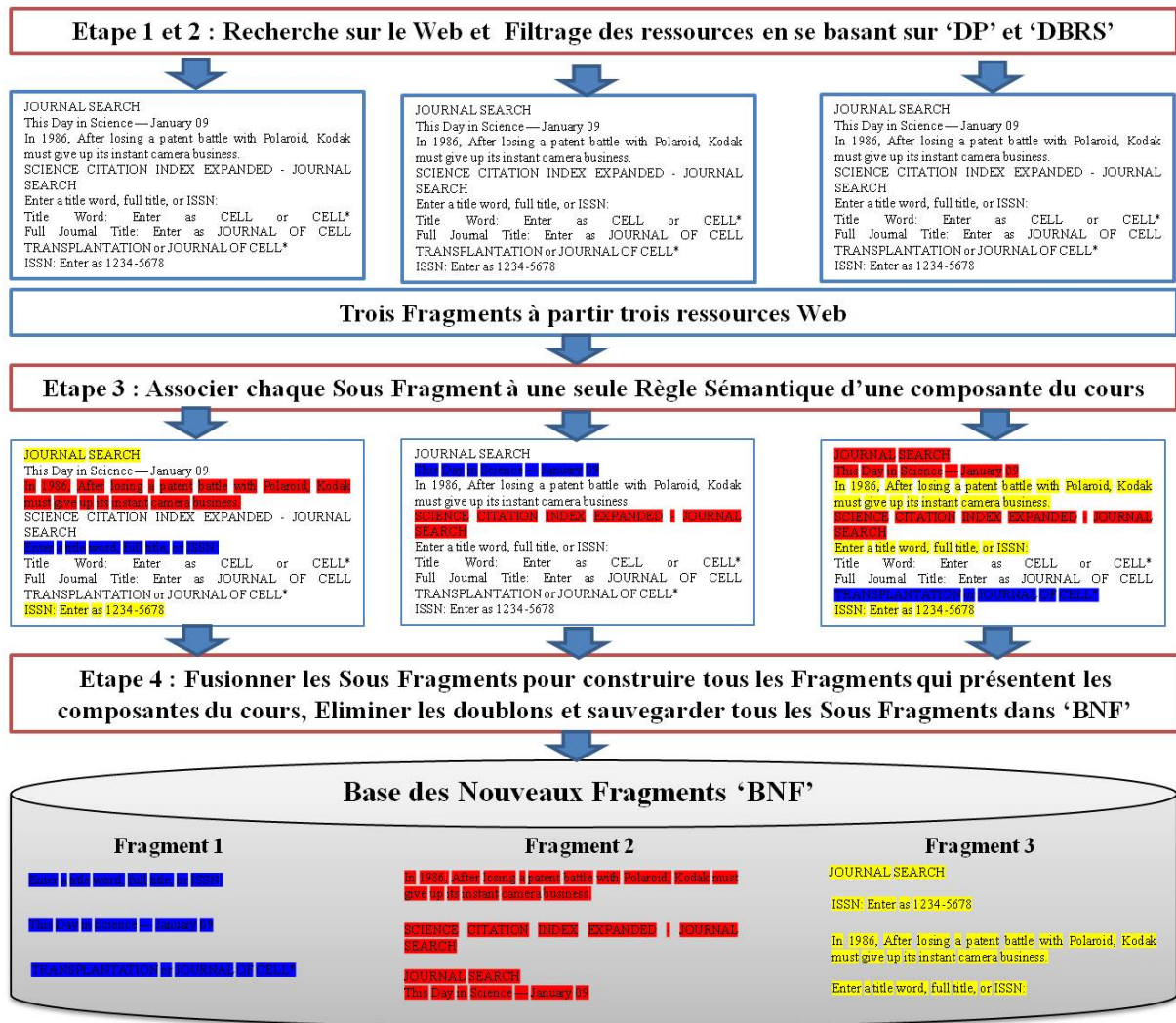


Figure 17 Le principe du processus de fusion de trois ressources Web

Dans cet exemple, nous avons choisi à partir la base 'BNF' les sous fragments qui ont été associés aux règles sémantiques. Il ya trois parties du cours avec trois couleurs différentes (bleu, rouge et jaune). Après cela, nous pouvons commencer la fusion pour chaque composante du cours et sauvegarder le résultat dans la case correspondante dans le modèle Excel 'FSF' (Voir Tableau 5 en page 133).

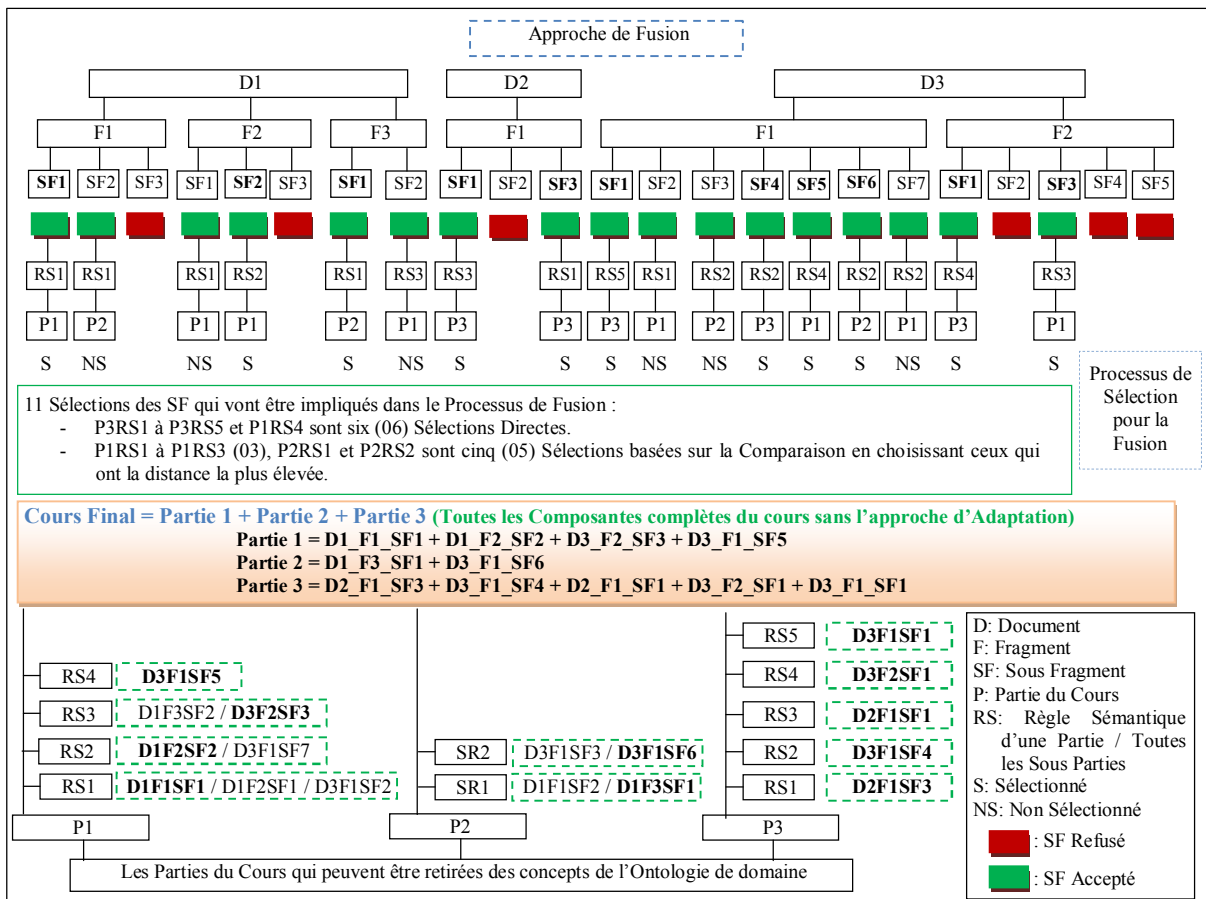


Figure 18 Fusion des sous fragments

8.3.3. L'approche de la Fusion des sous fragments

La Fusion est basée sur les règles sémantiques 'RS' entre les termes dans chaque sous partie du cours (Voir Figure 18 ci-dessus) où chaque sous partie est présentée par une règle sémantique. Donc, un concept de l'ontologie (composante du cours) est présenté par plusieurs 'RS'. Ces dernières sont créées pour toutes les composantes du cours. Chacune d'entre elles doit être associée à une seule composante du cours.

Dans notre approche, l'ontologie de domaine joue trois rôles :

- **Le premier**, est la présentation de toute la hiérarchie de la formation, comme nous l'avons mentionné dans le chapitre précédent (Voir le titre 7.9.2 en page 121), l'ontologie dans ce niveau, permet le calcul du degré de pertinence 'DP' (Voir Formule 1 en page 126 et Formule 2 en page 128).

- **Le second**, est la sauvegarde des règles sémantiques 'RS' qui permettent le calcul de la méthode 'DBRS' avec ses trois formules : $DBRS_M$ (Moyenne), $DBRS_D$ (Distance) et $DBRS_{PA}$ (Précision de l'association) (Voir Formule 3, Formule 4 et Formule 5 en page 129, 130 et 131 respectivement). La méthode est basée essentiellement sur les distances entre les termes des 'RS' par rapport aux mots des sous fragments.
- **Le troisième**, est d'offrir la possibilité de la fusion des sous fragments trouvés après la recherche, le filtrage et la sélection. Car, chaque concept de l'ontologie a un ensemble de 'RS' qui permettent la Fusion des sous fragments extraits à partir du Web qui sont associés aux 'RS'.

La fusion peut avoir lieu dans les deux cas suivants : la fusion d'une ou plusieurs ressources Web et la fusion entre les sous fragments de la même ressource Web. Dans ces deux cas, nous pouvons trouver des sous fragments en plusieurs emplacements dans la ressource Web. Il y'a deux niveaux différents d'extraction (Voir Figure 18 en page 138) : (1) : Le niveau d'un Fragment, et (2) : Le niveau d'un Sous Fragment.

8.3.3.1. Le niveau d'un Fragment

Il est basé sur la méthode 'DP', nous appliquons la méthode dans un premier temps sur toutes les ressources Web pour trouver celles les plus pertinentes par rapport au sujet de la recherche ou le cours complet. Dans la Figure 18 en page 138, nous avons trois ressources Web qui sont : 'D1', 'D2' et 'D3' qui présentent trois documents différents. Après cela, nous appliquons la méthode 'DP' pour une autre fois pour toutes les ressources Web les plus pertinentes pour extraire des fragments les plus pertinents par rapport aux composantes du cours comme dans la Figure 18 en page 138. Par exemple les Fragments 'F1, F2 & F3' sont les fragments les plus pertinents dans le document 'D1'. Ce dernier peut contenir d'autres fragments.

8.3.3.2. Le niveau d'un Sous Fragment

Il est basé sur la méthode 'DBSR' qui contient les trois formules : $DBRS_M$ (Moyenne), $DBRS_D$ (Distance) et $DBRS_{PA}$ (Précision de l'association) (Voir Formule 3, Formule 4 et Formule 5 en page 129, 130 et 131 respectivement), ces trois formules sont basés sur les mots des sous fragments des ressources Web, les mots sont comparés avec les termes des règles sémantiques 'RS'.

Nous appliquons la méthode sur tous les fragments extraits dans le premier niveau, pour extraire des sous fragments 'SF' les plus appropriés par rapport aux règles sémantiques 'RS'. Par la suite, nous associons les Sous Fragments de la couleur verte aux règles sémantiques correspondantes comme dans la Figure 18 en page 138. Par exemple les Sous Fragments 'SF1' et 'SF2' sont sélectionnés à partir des Sous Fragments 'SF1, SF2 et SF3' du Fragment 'F1' du Document 'D1'.

En conséquence, nous extrayons trois sous fragments: D1F1SF1 / D1F2SF1 / D3F1SF2 par rapport à la règle sémantique 'RS1' dans la partie 'P1' ; 'P1' est la première partie du cours qui comprend quatre règles sémantiques 'RS1, RS2, RS3 et RS4'. Pour la 'RS1' nous choisissons uniquement le Sous Fragment en gras 'D1F1SF1' des trois sous fragments associés à cette règle sémantique. 'D1F1SF1' : Sous Fragment un (01) du Fragment un (01) du Document un (01).

Après cela, nous sélectionnons tous les sous fragments qui sont mentionnés en gras dans la Figure 18 en page 138 pour obtenir les composantes de chaque partie du cours. Par exemple : la partie une (01) 'P1' contient quatre sous fragments sélectionnés à partir quatre fragments différents qui sont extraits à partir deux documents Web différents :

Partie 1 = D1_F1_SF1 + D1_F2_SF2 + D3_F2_SF3 + D3_F1_SF5 selon la Figure 18 en page 138 'Les deux Sous fragment : un (01) du fragment un (01) et deux (02) du fragment deux (02), du document un (01). Ainsi, les deux Sous fragment : trois (03) du fragment deux (02) et cinq (05) du fragment un (01), du document trois (03)'.

Les résultats de cette approche de Fusion vont être enregistrés dans la base 'BNF'. C'est-à-dire, chaque sous fragment pertinent doit être sauvegardé dans la ligne correspondante dans le modèle Excel (Voir Tableau 4 en page 133). Les sous fragments de la même ligne doivent être ordonnés selon la Formule 5 en page 131.

8.4. Construction du domaine d'apprentissage 'DA'

Après le processus de fusion (Voir Figure 18 en page 138), nous obtenons tous les sous fragments qui sont enregistrés dans la base 'BNF'. Pour cette raison, la base 'BNF' présente notre domaine d'apprentissage 'DA'. Le domaine 'DA' peut également être adapté aux profils des apprenants, parce que notre ontologie pédagogique repose sur les mêmes règles sémantiques qui sont définies dans l'ontologie de domaine 'Squelette Humain'.

Pour cet objectif nous associons chaque sous fragment extrait à partir des ressources Web à une seule règle sémantique qui permet la création d'une sous partie du cours. Nous pouvons réaliser l'approche d'adaptation facilement avec ce principe pour adapter le contenu de la base 'BNF' aux profils des apprenants.

8.5. L'Adaptation dans PrOALDoF-Web

L'Adaptation du contenu du système PrOALDoF-Web est un mécanisme qui permet d'adapter les briques d'informations 'sous fragments' stockées dans la base 'BNF' aux profils des apprenants et selon un contexte bien défini par les auteurs de la formation à suivre. Le contexte de la formation est défini dans l'ontologie de domaine qui contient toutes les composantes d'un cours (Voir Figure 13 en page 120).

Un profil d'apprenant est présenté uniquement par un niveau dans PrOALDoF-Web. Donc, les préférences et les besoins d'un apprenant sont exprimés par son niveau. Tous les niveaux de la formation sont stockés dans l'ontologie pédagogique (Voir Figure 19 ci-dessous) où chaque règle sémantique est associée à un seul niveau.

L'ontologie pédagogique est créée en ajoutant une référence du niveau 'N1' qui présente le niveau numéro un (01) aux références des règles sémantiques qui sont définies dans l'ontologie de domaine, pour que nous pouvons trouver directement la ligne correspondante dans le base 'BNF' dans l'approche d'Adaptation.

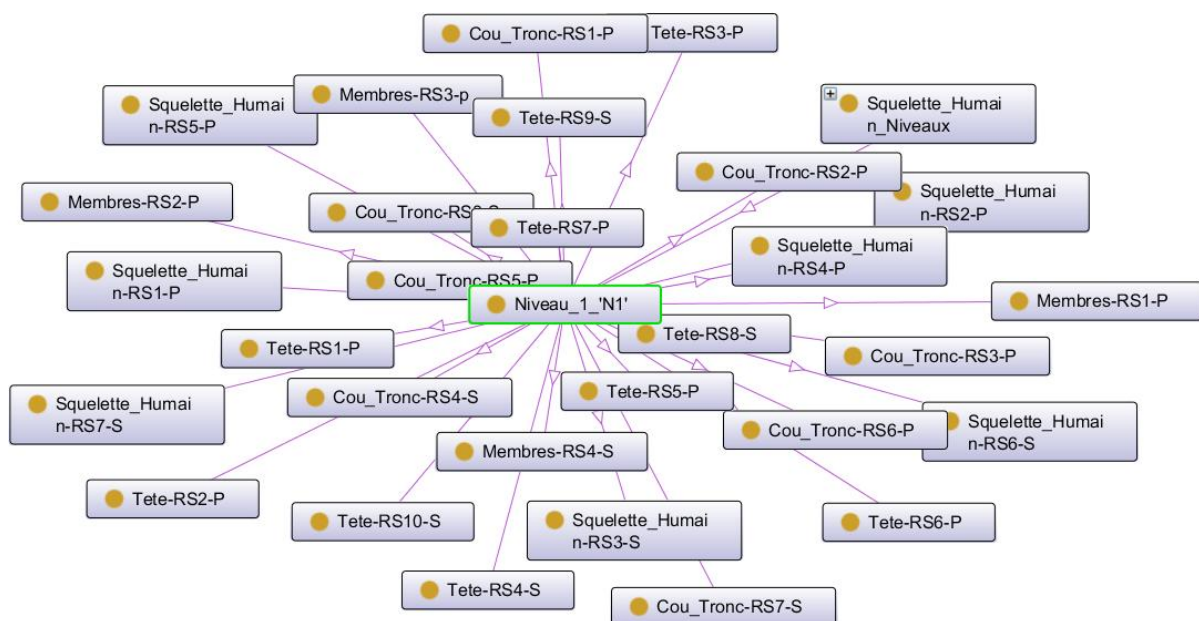


Figure 19 Extrait graphique de l'ontologie Pédagogique

Dans la Figure 19 en page 141, nous avons un extrait graphique des règles sémantiques de quatre concepts de l'ontologie de domaine. Les règles sont associées au niveau numéro un (01) dans l'ontologie pédagogique.

Chaque règle sémantique qui présente une ligne dans le modèle Excel, est définie dans l'ontologie de domaine comme suit :

C1-...-CN-RSN°

Où

C1-...-CN : représente l'arborescente de la règle sémantique 'RS' qui a le numéro 'N°'.

Exemple : Squelette_Humain-Membres-Membres_Inférieurs-Main-RS7

Où

'RS7' est la septième règle sémantique du concept 'Main'.

Donc, dans l'ontologie pédagogique, nous gardons la définition de chaque règle sémantique dans l'ontologie de domaine. Nous obtenons le niveau de chaque règle à partir la hiérarchie des niveaux dans l'ontologie pédagogique comme suit (Voir Figure 20 ci-dessous) :

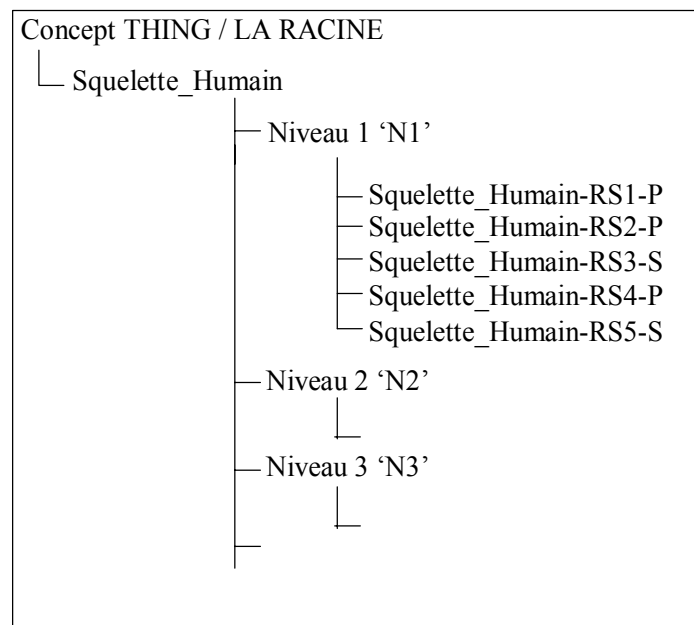


Figure 20 Extrait graphique de la hiérarchie de l'ontologie Pédagogique

Selon l'extrait de la hiérarchie (Voir Figure 20 ci-dessus), nous avons quelques règles sémantiques qui sont définies pour le concept 'Squelette_Humain'.

- **Exemples**

‘**Squelette_Humain-RS1-P**’ : est la première règle sémantique qui présente un pré-requis pour la composante du cours ‘Squelette Humain’.

Squelette_Humain-RS3-S : est la troisième règle sémantique qui présente une information supplémentaire pour la même composante du cours.

L'Adaptation est basée sur le principe du pré-requis comme relation entre les composantes de base d'un cours. Ainsi, sur le principe des informations supplémentaires en donnant d'autres informations pour certaines composantes du cours.

Nous présentons ces deux principes dans l'ontologie pédagogique comme étant un graphe où le concept père présente le niveau le plus bas, chaque relation à partir de ce concept présente une règle sémantique. Il y a deux types de règles : pré-requis et information supplémentaire.

Toutes les règles sémantiques d'un niveau sont dans le même niveau dans l'ontologie pédagogique. Donc, Nous distinguons la différence entre un pré-requis et une information supplémentaire par l'objectif de la formation en rajoutant une autre information dans la fin de la référence d'une règle sémantique comme suit :

C1-...-CN-RSN^o-NN^o-P ou C1-...-CN-RSN^o-NN^o-S

Où

P : représente un pré-requis ;

S : représente une information supplémentaire.

En conséquence, l'ontologie pédagogique contient tout le processus de navigation dans le système via les différents niveaux définis et toutes les relations entre les composantes du cours et les règles sémantiques associées aux niveaux.

8.6. Un domaine d'apprentissage adaptatif 'DAA'

Notre méthode d'adaptation est basée sur les niveaux des apprenants. Grâce à cette méthode, nous pouvons adapter les cours aux apprenants. Nous n'avons pas travaillé sur le modèle apprenant, car notre but est la génération du cours automatiquement via les ressources Web aux profils des apprenants. Donc, nous supposons que nous avons les niveaux dès le début, et chaque navigation dans le projet PrOALDoF-Web augmente le niveau de l'apprenant en acquérant d'autres composantes du cours.

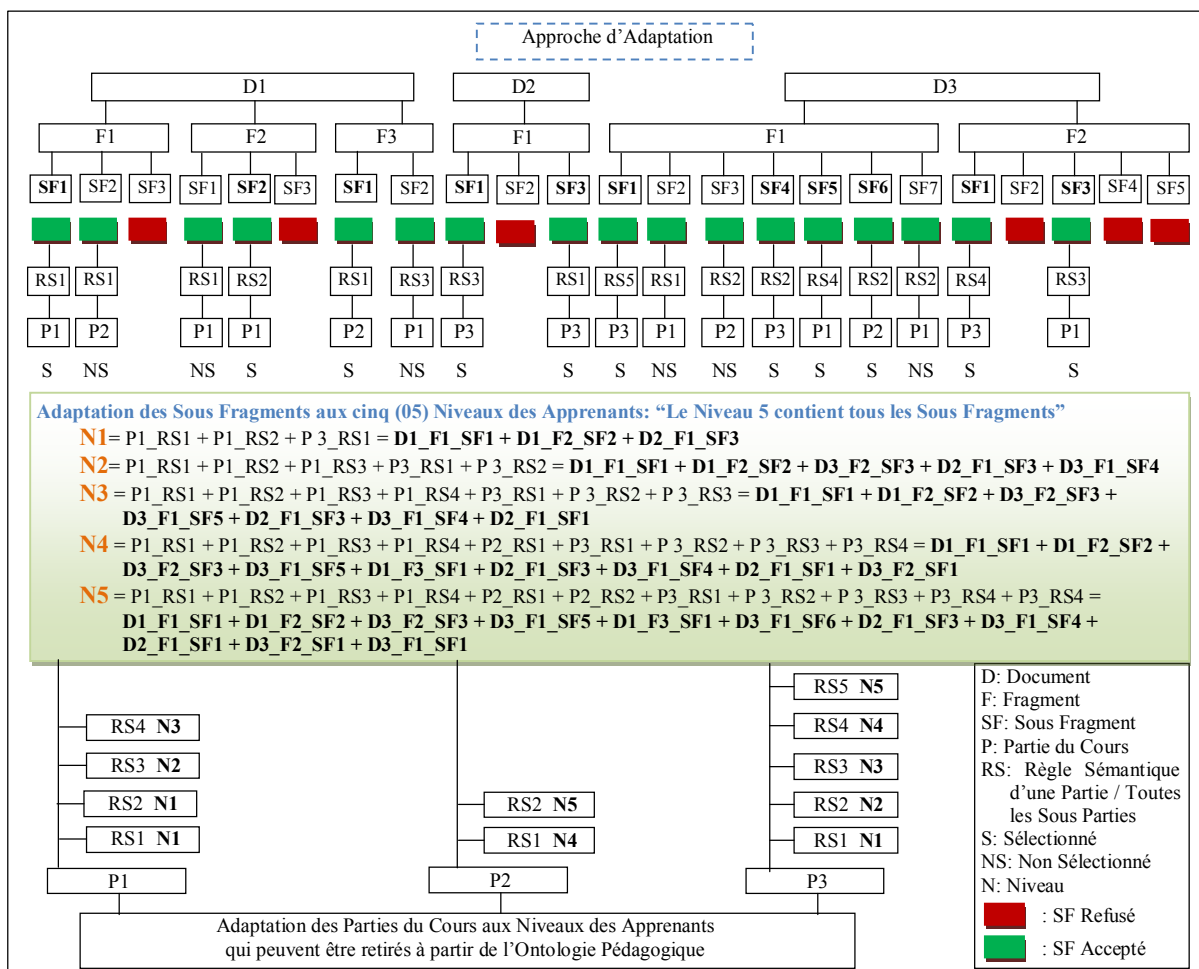


Figure 21 L'approche d'Adaptation basée sur la fusion des sous fragments

Nous pouvons adapter les cours grâce à notre ontologie pédagogique car nous avons le niveau de l'apprenant. Dans notre approche, nous avons créé l'ontologie pédagogique avec les mêmes règles sémantiques définies dans l'ontologie de domaine, mais les 'RS' sont regroupées en niveaux. C'est-à-dire, chaque niveau de l'apprenant a un certain nombre de composantes du cours. Chacune d'elles a un ensemble de règles sémantiques. Lorsque le niveau de l'apprenant augmente, le nombre de 'RS' augmente aussi (Voir Figure 21 ci-dessus). Chaque règle sémantique représente une sous partie du cours.

A la fin de notre travail, nous allons atteindre la réalisation du domaine d'apprentissage adaptatif 'DAA' à partir des ressources Web. Nous pouvons alors, détailler le processus d'Adaptation selon la Figure 21 ci-dessus. Nous avons associé chaque règle sémantique 'RS' à un (01) seul niveau 'N' (RS1 N1: Règle Sémantique numéro un (01) est associée au niveau numéro un (01)). Pour chaque partie du cours 'P', nous devons spécifier toutes les règles sémantiques lui associées ('P1' est la première partie du cours qui contient quatre (04) règles sémantiques : 'RS1 N1, RS2 N1, RS3 N2 et RS4 N3').

Après l'étape d'extraction des sous fragments les plus pertinents, nous associons chaque sous fragment 'SF' à une seule règle sémantique d'une partie du cours 'P'. Avec la fusion des règles sémantiques d'un ou plusieurs niveaux, nous pouvons savoir quels sont les sous fragments du niveau de l'apprenant présent qui doivent être sélectionnés.

Le niveau numéro un (01) est composé de deux parties du cours 'P1' et 'P3', avec trois 'RS' : 'P1_RS1', 'P1_RS2' et 'P3_RS1'. Les sous fragments les plus pertinents pour ces trois 'RS' sont comme suit : 'D1_F1_SF1', 'D1_F2_SF2' et 'D2_F1_SF3' où 'D1_F1_SF1' est le premier sous fragment du premier fragment du premier document Web. 'D1_F2_SF2' est le second sous fragment du deuxième fragment du premier document. Enfin, 'D2_F1_SF3' est le troisième sous fragment du premier fragment du deuxième document. L'extraction des sous fragments est basée sur le processus de fusion.

Avec ce principe, nous obtenons tous les sous fragments de chaque niveau. Nous faisons la fusion de ces derniers afin de créer un cours adaptatif à l'apprenant.

8.7. Conclusion

Le PrOALDoF-Web a fait l'apparition d'un système éducatif en ligne capable de trouver tous les éléments et toutes les ressources indispensables au fonctionnement du processus d'enseignement d'un apprenant. Il vise à trouver dans un premier lieu tous les documents pertinents par rapport au sujet de recherche. Dans un deuxième lieu, il filtre les documents pour extraire les fragments les plus pertinents par rapport aux composantes du cours.

Ensuite, il extrait les sous fragments appropriés aux règles sémantiques définies pour chaque composante du cours. Il a une approche de fusion, donc, il peut fusionner les sous fragments pour créer de nouveaux fragments. Enfin, il peut adapter les sous fragments aux profils des apprenants, en appuyant sur les mêmes règles sémantiques associées aux composantes du cours, mais dans ce niveau, elles sont associées aux niveaux de la formation. Les règles sont définies et sauvegardées dans l'ontologie pédagogique.

Le chapitre suivant présente nos résultats obtenus avec une évaluation du projet PrOALDoF-Web en comparaison avec les différentes méthodes de création des cours.

9. Evaluation du PrOALDoF-Web

9.1. Introduction

A travers ce chapitre, nous présentons dans un premier lieu un exemple illustrant tout le processus essentiel dans notre projet qui est la fusion et l'adaptation. Car la recherche est un processus à l'intérieur du système que nous ne pouvons pas le visualisé. Nous avons le simulé via la recherche manuelle avec le moteur de recherche Google présentée dans le chapitre II dans la deuxième partie de ce document en page 111. Par la suite, nous comparons 'PrOALDoF-Web' avec les outils de création des cours existants. Ces derniers ont deux types : commercial ou éducatif.

Le but de l'évaluation est de mettre en position nos résultats et de montrer la performance de notre système proposé. Car nous avons une comparaison entre un processus automatique, et un autre manuel ou humain réalisé par les auteurs. Chaque outil de création des cours nécessite une forte concentration humaine en temps et en effort.

Bien sur, nous jugeons la qualité dans ce présent chapitre par le fait que les règles sémantiques qui sont impliquées dans la tâche d'extraction des sous fragments les plus pertinents qui sont appropriés aux besoins de la formation, sont créées par des experts dans le domaine. Ainsi, via notre ontologie de domaine qui a des annotations et des mots clés impliqués dans la recherche Web et le filtrage des résultats trouvés pour extraire les fragments les plus pertinents, l'ontologie de domaine est créée en utilisant l'expertise des experts du domaine.

Le chapitre montre l'efficacité de notre approche proposée en présentant tout le processus de fusion et d'adaptation via un exemple illustratif. Ceci se voit par le fait que l'approche trouve les éléments nécessaires qui sont appropriés aux besoins de la formation et surtout que toute l'approche fonctionne d'une façon automatique.

Une deuxième évaluation sera détaillée plus tard, se focalise sur l'utilisabilité du système, que ce soit par les apprenants qui suivent leurs cours ou bien par les auteurs qui examinent le contenu à chaque création automatique pour qu'ils jugent la qualité du contenu. Un apprenant peut nous aider à améliorer la façon de présenter le cours, car celui qui juge la qualité est bien entendu l'auteur qui est expert dans sa matière.

9.2. Exemple illustratif

Nous illustrons à partir cet exemple les différentes méthodes impliquées dans l'approche de fusion. Puis, nous montrons la façon d'adaptation du cours aux apprenants via leurs niveaux.

Les données d'entrées sont comme suit :

Neuf (09) sous fragments 'SF' :

SF1 = Le **squelette** d'un adulte comporte **206 os**, il est une partie du **corps** humain et protège les **organes** internes.

SF2 = Le **squelette** d'un être humain comporte **206 os**, il fait partie du **corps** humain.

SF3 = Le **squelette** constitue la structure du **corps** humain et protège les **organes** internes.

SF4 = Le **corps** humain est composé d'un système squelettique (**squelette**), et d'un ensemble d'**organes**.

SF5 = Le **squelette** humain est composé de **206 os** constants.

SF6 = Le **squelette** est la charpente de l'anatomie humaine, qui soutient le **corps** et protège ses **organes** internes.

SF7 = A la naissance, les **bébés** ont environ **350 os**, dont une grande partie va se souder au cours de la croissance.

SF8 = A l'adulte, le **squelette** est composé de **206 os**.

SF9 = Le **squelette** fait partie du **corps** humain, il protège les **organes**, un **bébé** possède à sa naissance en moyenne **350 os**, la plupart de ces os fusionnent quand il grandie.

Trois (03) règles sémantiques 'RS' :

RS1 = Squelette-corps-organes.

RS2 = Squelette-206 os.

RS3 = Bébé-350 os.

Sept (07) termes extraits à partir des trois (03) règles RS1, RS2 et RS3 :

RS1_T1 = Squelette ;

RS1_T2 = Corps ;

RS1_T3 = Organes ;

RS2_T1 = Squelette ;

RS2_T2 = 206 os ;

RS3_T1 = bébé ;

RS3_T2 = 350 os.

Deux (02) niveaux ‘N1’ et ‘N2’ :

RS1_N1 = la première règle sémantique est associée au niveau 1 ;

RS2_N1 = la deuxième règle sémantique est associée au niveau 1 ;

RS3_N2 = la troisième règle sémantique est associée au niveau 2.

Nous avons trois parties dans les calculs et les traitements : Calcul des Moyennes, calcul des Distances et calcul des Précisions d’Associations. Puis, nous avons l’application de la méthode de fusion globale dans la quatrième partie pour créer le cours complet. Ainsi, l’application de la méthode d’adaptation dans la cinquième partie pour adapter un cours à un apprenant particulier.

9.2.1. Partie 1

Le calcul de la moyenne (Formule 3 en page 129) des existences des termes **RS1_T1**, **RS1_T2**, **RS1_T3**, **RS2_T1**, **RS2_T2**, **RS3_T1** et **RS3_T2** dans les sous fragments **SF1**, **SF2**, **SF3**, **SF4**, **SF5**, **SF6**, **SF7**, **SF8** et **SF9** en se basant sur les trois règles sémantiques **RS1**, **RS2** et **RS3**. La lettre ‘**M**’ désigne le marquage d’un mot dans un sous fragment. Nous effectuons ce marquage pour chaque sous fragment.

9.2.1.1. SF1 par rapport RS1, RS2 et RS3

SF1_RS1 = Le *squelette* (**M1**) d’un *adulte* (**M2**) *comporte* (**M3**) 206 (**M4**) *os* (**M5**), il est une *partie* (**M6**) du *corps* (**M7**) *humain* (**M8**) et *protège* (**M9**) les *organes* (**M10**) *internes* (**M11**).

$$DBRS_M = (1+1+1)/11 = 3/11 = 0,27$$

SF1_RS2 = Le *squelette* (**M1**) d’un *adulte* (**M2**) *comporte* (**M3**) **206** (**M4**) *os* (**M5**), il est une *partie* (**M6**) du *corps* (**M7**) *humain* (**M8**) et *protège* (**M9**) les *organes* (**M10**) *internes* (**M11**).

$$DBRS_M = (1+1)/11 = 2/11 = 0,18$$

SF1_RS3 = Le *squelette* (**M1**) d’un *adulte* (**M2**) *comporte* (**M3**) 206 (**M4**) *os* (**M5**), il est une *partie* (**M6**) du *corps* (**M7**) *humain* (**M8**) et *protège* (**M9**) les *organes* (**M10**) *internes* (**M11**).

$$DBRS_M = (0+0)/11 = 0/11 = 0$$

9.2.1.2. SF2 par rapport RS1, RS2 et RS3

SF2_RS1 = Le *squelette* (M1) d'un être (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

$$DBRS_M = (1+1+0)/10 = 2/10 = 0,2$$

SF2_RS2 = Le *squelette* (M1) d'un être (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

$$DBRS_M = (1+1)/10 = 2/10 = 0,2$$

SF2_RS3 = Le *squelette* (M1) d'un être (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

$$DBRS_M = (0+0)/10 = 0/10 = 0$$

Nous continuons avec le même principe pour tous les sous fragments par rapport aux règles sémantiques. Nous obtenons à la fin toutes les moyennes avec un marquage des termes trouvés dans chaque sous fragments (Voir Tableau 7 en page 152).

9.2.2. Partie 2

Le calcul de la distance (Formule 4 en page 130) entre les termes **RS1_T1**, **RS1_T2**, **RS1_T3**, **RS2_T1**, **RS2_T2**, **RS3_T1** et **RS3_T2** existants dans les sous fragments **SF1**, **SF2**, **SF3**, **SF4**, **SF5**, **SF6**, **SF7**, **SF8** et **SF9** en se basant sur les marquages 'M' effectués dans le calcul des moyennes dans la première partie.

9.2.2.1. SF1 par rapport RS1, RS2 et RS3 :

SF1_RS1 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) 206 (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

$$DBRS_D = (7-1+10-7) = 6+3 = 9$$

SF1_RS2 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) 206 (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

$$DBRS_D = (4,5-1) = 3,5$$

SF1_RS3 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) 206 (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

Marquage non complet. Donc, pas de calcul de la distance pour **SF1_RS3**.

9.2.2.2. SF2 par rapport RS1, RS2 et RS3

SF2_RS1 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

Marquage non complet. Donc, pas de calcul de la distance pour **SF2_RS1**.

SF2_RS2 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) **206** (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

$$DBRS_D = (5,5-1) = 4,5$$

SF2_RS3 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

Marquage non complet. Donc, pas de calcul de la distance pour **SF2_RS3**.

Nous continuons avec le même principe pour tous les sous fragments qui ont les termes des règles sémantiques. Nous obtenons à la fin toutes les distances dans tous les sous fragments (Voir Tableau 7 en page 152).

9.2.3. Partie 3

Le calcul de la précision d'association (Formule 5 en page 131) des sous fragments **SF1**, **SF2**, **SF3**, **SF4**, **SF5**, **SF6**, **SF7**, **SF8** et **SF9** par rapport aux règles sémantiques **RS1**, **RS2** et **RS3** en se basant sur les marquages 'M' effectués dans le calcul des moyennes dans la première partie. Le poids 'P' égal à 0,7.

9.2.3.1. SF1 par rapport RS1, RS2 et RS3

SF1_RS1 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) 206 (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

$$DBRS_M = 0,27, DBRS_D = 9 \text{ et } NB_{te} = 3$$

$$DBRS_{PA} = 0,27 * \frac{3}{9} = 0,09$$

SF1_RS2 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) **206** (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

$DBRS_M = 0,18$, $DBRS_D = 3,5$ ET $NB_{te} = 2$

$$DBRS_{PA} = 0,18 * \frac{2}{3,5} = 0,10$$

SF1_RS3 = Le *squelette* (M1) d'un *adulte* (M2) *comporte* (M3) 206 (M4) *os* (M5), il est une *partie* (M6) du *corps* (M7) *humain* (M8) et *protège* (M9) les *organes* (M10) *internes* (M11).

Marquage non complet. Donc, pas de calcul de la précision d'association pour le sous fragment **SF1_RS3**.

9.2.3.2. SF2 par rapport RS1, RS2 et RS3

SF2_RS1 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

Marquage non complet. Donc, pas de calcul de la précision d'association pour le sous fragment **SF2_RS1**.

SF2_RS2 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) **206** (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

$DBRS_M = 0,2$, $DBRS_D = 4,5$ ET $NB_{te} = 2$

$$DBRS_{PA} = 0,2 * \frac{2}{4,5} = 0,09$$

SF2_RS3 = Le *squelette* (M1) d'un *être* (M2) *humain* (M3) *comporte* (M4) 206 (M5) *os* (M6), il *fait* (M7) *partie* (M8) du *corps* (M9) *humain* (M10).

Marquage non complet. Donc, pas de calcul de la précision d'association pour le sous fragment **SF2_RS3**.

Nous continuons avec le même principe pour tous les sous fragments qui satisfont l'existence des termes des règles sémantiques. Nous obtenons à la fin toutes les précisions d'associations de tous les sous fragments par rapport aux règles sémantiques (Voir Tableau 7 en page 152).

	Nb Mots	Existence	Marquage	DBRS _M	Sélection	DBRS _D	DBRS _{PA}
SF1_RS1	11	3	M1, M7, M10	0,27	Oui	9	0,09
SF1-RS2	11	2	M1, (M4, M5)	0,18	Oui	3,5	0,10
SF1-RS3	11	0	/	0,00	Non	/	/
SF2_RS1	10	2	M1, M9	0,20	Non	/	/
SF2-RS2	10	2	M1, (M5, M6)	0,20	Oui	4,5	0,09
SF2-RS3	10	0	/	0,00	Non	/	/
SF3_RS1	8	3	M1, M4, M7	0,38	Oui	6	0,19
SF3-RS2	8	1	M1	0,13	Non	/	/
SF3-RS3	8	0	/	0,00	Non	/	/
SF4_RS1	8	3	M1, M6, M8	0,38	Oui	7	0,16
SF4-RS2	8	1	M6	0,13	Non	/	/
SF4-RS3	8	0	/	0,00	Non	/	/
SF5_RS1	6	1	M1	0,17	Non	/	/
SF5-RS2	7	2	M1, (M4, M5)	0,28	Oui	3,5	0,16
SF5-RS3	6	0	/	0,00	Non	/	/
SF6_RS1	9	3	M1, M6, M8	0,33	Oui	7	0,14
SF6-RS2	9	1	M1	0,11	Non	/	/
SF6-RS3	9	0	/	0,00	Non	/	/
SF7_RS1	11	0	/	0,00	Non	/	/
SF7-RS2	11	0	/	0,00	Non	/	/
SF7-RS3	11	2	M2, (M5, M6)	0,18	Oui	3,5	0,10
SF8_RS1	6	1	M2	0,17	Non	/	/
SF8-RS2	6	2	M2, (M5, M6)	0,33	Oui	3,5	0,19
SF8-RS3	6	0	/	0,00	Non	/	/
SF9_RS1	18	3	M1, M4, M7	0,17	Oui	6	0,09
SF9-RS2	18	1	M1	0,06	Non	/	/
SF9-RS3	18	2	M8, (M12, M13)	0,11	Oui	4,5	0,05

Tableau 7 Les résultats obtenus en appliquant la méthode ‘DBRS’

Si nous trouvons deux marquages entre parenthèse, c'est-à-dire, c'est un mot d'un terme composé de plusieurs termes tel que : (M4, M5) dans SF1-RS2, le terme est : 206 os. Donc, la distance dans ce cas se calcule en faisant la moyenne des termes. Le nouveau marquage est : 4,5. La distance égale à 3,5.

S'il y a une égalité dans la 'DBRS_{PA}', nous prenons le sous fragment qui a le nombre de mots le plus petit, pour assurer qu'il n'y a pas des répétitions par la suite dans les méthodes de fusion et d'adaptation.

9.2.4. Partie 4

La Fusion du cours complet des sous fragments les plus pertinents SF3_RS1, SF8-RS2, et SF7-RS3 pour construire cette partie du cours avec ces deux niveaux 'N1' et 'N2'. Le résultat de la fusion sera comme suit :

Le **squelette** constitue la structure du **corps** humain et protège les **organes** internes. A l'adulte, le **squelette** est composé de **206 os**. A la naissance, les **bébés** ont environ **350 os**, dont une grande partie va se souder au cours de la croissance.

Nous remarquons par exemple pour la troisième règle sémantique 'RS3' qu'il y a deux sous fragments qui contiennent ses termes 'SF7' et 'SF9'. Le premier contient 6 mots et le second contient 18 mots. Donc nous voyons déjà une grande différence en nombre de mots. Intuitivement, nous pensons que le meilleur sous fragment est 'SF7'. Par calcul, nous trouvons les précisions de l'association respectivement comme suit : 0,10 et 0,05. Donc, le meilleur sous fragment pour la règle sémantique 'RS3' est 'SF7'.

9.2.5. Partie 5

L'Adaptation du cours au niveau 'N1' qui va être avec les deux sous fragments les plus pertinents SF3_RS1 et SF8-RS2. Le résultat de l'Adaptation sera comme suit :

Le **squelette** constitue la structure du **corps** humain et protège les **organes** internes. A l'adulte, le **squelette** est composé de **206 os**.

Nous remarquons qu'il existe uniquement deux niveaux dans cet exemple. Donc, l'adaptation du cours au niveau 'N2' va être avec les trois règles sémantiques associées à ce niveau dont le contenu de ce dernier présente le cours complet défini dans la quatrième partie 'Fusion'.

9.3. Comparaison entre PrOALDoF-Web et les outils de création des cours

Nous avons deux points dans cette section : les éléments de la comparaison et les résultats de cette comparaison.

9.3.1. Les éléments de la comparaison

Nous avons 14 points ou éléments pour évaluer la qualité et l'intérêt pour utiliser notre approche, afin d'obtenir plus d'avantages que les autres méthodes de création des cours et bien entendu les systèmes d'enseignement en ligne.

Nous avons comparé PrOALDoF-Web avec trois méthodes de création des cours :

- Création manuelle par les auteurs ;
- Création semi-automatique des cours dans les plateformes d'E-Learning, cette méthode est assistée car les auteurs utilisent des outils de création ;
- Enfin, création dynamique, elle est aussi semi-automatique, car les cours sont préalablement créés dans le support physique de la formation. Nous trouvons cette création dans les recherches récentes dans les plateformes d'enseignement.

Nous avons obtenu 9 points à partir l'instrument d'évaluation des objets d'apprentissage 'Learning Object Review Instrument' (LORI 1.5) qui est un Framework d'évaluation des ressources d'apprentissage multimédia (Nesbit, Belfer et Leacock, 2004) (Voir Tableau 8 ci-dessous).

Item	Brief Description
Content quality	Veracity, accuracy, balanced presentation of ideas, and appropriate level of detail
Learning goal alignment	Alignment among learning goals, activities, assessments, and learner characteristics
Feedback and adaptation	Adaptive content or feedback driven by differential learner input or learner modeling
Motivation	Ability to motivate and interest an identified population of learners
Presentation design	Design of visual and auditory information for enhanced learning and efficient mental processing
Interaction usability	Ease of navigation, predictability of the user interface, and the quality of the interface help features
Accessibility	Design of controls and presentation formats to accommodate disabled and mobile learners
Reusability	Ability to use in varying learning contexts and with learners from different backgrounds
Standards compliance	Adherence to international technical standards and specifications

Tableau 8 Eléments de comparaison dans LORI 1.5

En dessous des ces éléments, cinq autres éléments sont proposés dans le processus d'évaluation de PrOALDoF-Web (Chaoui et Laskri, 2013) (Voir Tableau 9 ci-dessous).

Item	Brief Description
Method of creation	Manual or Semi-Automatic or Automatic way of courses' creation
Enrichment method	Process of adding further information in courses
Ease of Use	Critical to the last user and the customer. It includes not only aspects such as how the information and suggestions are presented to the last user, but also the available documentation, the quality of it, and the effort required to learn how to use the system.
Content Research	Relevance of information to the domain of course and the quality of research
Experience of Authors	Authors at early stages of researches may face enormous difficulties of query formulation and how to decide the item or the part which must be chosen in order to create their own courses.

Tableau 9 Eléments de comparaison proposés dans PrOALDoF-Web

9.3.2. Les méthodes de création des cours

Dans notre étude, nous avons trois différentes méthodes de création des cours 'Manuelle', 'Semi-automatique' et 'Automatique'.

9.3.2.1. Méthode Manuelle

Elle est utilisée directement par les auteurs, ils cherchent les informations dans un domaine bien précis, ils utilisent différentes méthodes de recherche ou bien un ou plusieurs moteurs de recherche pour choisir les ressources d'information les plus pertinentes (Internet, livres, papiers, ...etc.). Les auteurs doivent filtrer les résultats afin de créer leurs propres cours en format électronique (Word, PDF ou autre format).

9.3.2.2. Méthode Semi-automatique

Elle contient deux sous-catégories : Méthode Assistée et Méthode Dynamique.

9.3.2.2.1. Méthode Assistée

C'est une méthode qui peut être assistée par un auteur, à l'aide des outils de la création des cours tels que : LMS 'Learning Management System' ou LCMS 'Learning Content Management System' qui sont disponibles sur les plateformes de formation comme : (Moodle, Sakai, dotLRN, OLAT, Olat LMS, OLMS, ...etc.).

Il y a aussi d'autres outils de création des cours qui présentent des logiciels gratuits tels que : (Adobe Coursebuilder, CALI Author, CourseLab, eXe, AuthorLCDS, Learning Content Development System, Lesson Writer, Memorize, MOS Solo ...etc.).

Enfin, il y a des outils de création des cours qui présentent des logiciels commerciaux tels que : (Adobe Captivate, Articulate, Camtasia, Compositica, Outstart Trainer, OpenWorld Presenter, PowerTrainer Authoring Tool ...etc.). Ces logiciels sont disponibles sur le marché.

Dans tous les types d'outils de création des cours cités précédemment, les auteurs doivent chercher l'information nécessaire à la création des cours, filtrer tous les résultats et extraire les informations les plus pertinentes pour suivre les indications de chacun des outils mentionnés ci-dessus pour terminer le processus de la création dans les bonnes conditions.

9.3.2.2.2. Méthode Dynamique

Selon l'aspect pédagogique qui sert à formuler des briques d'information. Bien sur, ces dernières sont déjà créées par des auteurs à travers le modèle proposé dans le système de formation en ligne. Nous parlons des systèmes qui proposent l'adaptation. Donc, le but est d'atteindre l'objectif de l'adaptation des cours aux profils des apprenants. Nous pouvons citer des exemples qui peuvent illustrer les travaux de recherche récents, à titre d'exemple les approches proposées en se basant sur : les ontologies, les systèmes multi-agents, ...etc.

9.3.2.3. Méthode Automatique

Grâce au système PrOALDoF-Web qui offre une approche qui permet d'éliminer le manque dans les deux premiers cas (méthode manuelle et semi-automatique). Nous pouvons localiser ce manque dans la recherche et le filtrage des informations nécessaires pour la création des cours.

Par ailleurs, l'approche fournit une occasion d'accroître la qualité des cours en combinant différentes ressources Web. Ainsi, elle donne un avantage d'adaptation des contenus aux profils des apprenants selon l'approche d'adaptation proposée pour les objectifs fixés de la formation.

9.3.3. Les résultats de la comparaison

Nous évaluons l'approche proposée en se basant sur les 14 points cités précédemment. L'évaluation est liée aux niveaux de comparaison qui sont détaillés dans le Tableau 10 en page 157 (Chaoui et Laskri, 2013). Chaque élément de la comparaison a une note entre 1 et 5. Il y a deux types de niveau : (1) : la manière de l'élément et (2) : la mesure de l'élément.

La manière d'un élément est caractérisée par le type de la méthode désigné dans l'élément. Tel que, le type de la méthode de création qui doit être comprise entre Manuelle '1 point' et Automatique '5 points'.

La mesure d'un élément se voit dans le traitement de l'outil de création des cours. Tel que, la facilité de l'utilisation de l'outil qui doit être comprise entre Faible '1 point' et Fort '5 point'.

Points of Comparison	1	2	3	4	5
Method of creation, Enrichment method, Content Research	Manual	Under Medium	Medium	Over Medium	Automatic
Ease of Use, Experience of Authors, Content quality, Feedback and adaptation, Learning goal alignment, Motivation, Presentation design, Interaction usability, Accessibility, Reusability, Standards Compliance	Low	Under Medium	Medium	Over Medium	High

Tableau 10 Les niveaux de comparaison

Le Tableau 11 ci-dessous montre une comparaison entre PrOALDoF-Web qui présente un outil de création automatique des cours et les différentes méthodes de création des cours. Ces dernières sont regroupées en deux catégories : la création manuelle par les auteurs, nous avons trois différents cours déjà créés qui entrent dans la comparaison 'C1, C2 et C3 : trois cours différents de trois auteurs différents'. La deuxième catégorie est la création semi-automatique qui a deux voies : la voie commerciale (MOODLE, ...etc.) et la voie académique via les dernières recherche dans le sens de la création des cours.

Points of Comparison \ Tool of Creation	Our Approach PrOALDoF-Web (Automatic way)	Manual creation (By Authors)			Creation of Course in E-Learning platforms (By Authors with tools' creation)	
		C1	C2	C3	Commercialized Platforms (MOODLE ...Etc.)	Research Platforms
Method of creation	5	1			1	3
Enrichment method	5	1			1	3
Ease of Use	5	1			1	5
Content Research	5	1			1	1
Experience of Authors	5	1	4	3	3	3
Content quality	5	2	5	2	3	4
Learning goal alignment	5	1			3	5
Feedback and adaptation	5	1			1	5
Motivation	5	1			2	5
Presentation design	3	1			3	4
Interaction usability	5	1			3	4
Accessibility	3	1			1	3
Reusability	5	1			1	4
Standards compliance	5	1			5	5

Tableau 11 Résultats de la comparaison

Nous indiquons les résultats du Tableau 11 en page 157 (Chaoui et Laskri, 2013) par des chiffres de 1 à 5 (Voir Tableau 10 en page 157).

Les résultats de la comparaison sont regroupés en 14 points que nous discutons chacun d'entre eux dans la section suivante.

9.3.3.1. Méthode de création

PrOALDoF-Web représente une méthode automatique de création des cours avec le Web comme source d'information. Le mécanisme de création est basé sur des méthodes statistiques et sémantiques 'DP' et 'DBRS'. Les trois étapes suivantes : la recherche, la fusion et l'adaptation, fonctionnent automatiquement.

Les autres outils de création nécessitent une recherche, un filtrage et une extraction des informations d'une façon manuelle. Un auteur veut créer son propre cours, doit préparer tout le processus de création. S'il a un outil de création des cours, il va extraire et mettre les briques d'informations demandées dans les emplacements appropriés. Il y a des chercheurs qui travaillent dans ce sens, mais la création reste toujours chez eux un processus qui demande une intervention humaine.

9.3.3.2. Méthode d'enrichissement

Dans PrOALDoF-Web il y a ce que nous avons appelé la base 'BNF'. Cette base contient tous les sous fragments jugés pertinents pour toutes les règles sémantiques définies dans le système. Ces dernières sont des sous parties des composantes des cours. L'approche propose une alternative mise à jour pour trouver des nouveaux sous fragments afin d'enrichir convenablement le contenu de la formation en ligne. Donc, même l'enrichissement dans PrOALDoF-Web se fait d'une façon automatique.

Le reste des outils exigent un effort personnel pour tout changement ou enrichissement des cours, car c'est le seul moyen pour mettre à jour les cours avec de nouvelles informations. Même chez les chercheurs qui travaillent sur la création dynamique des cours, ils ne peuvent pas enrichir le cours d'une façon automatique, mais, ils proposent des solutions pour créer un système dynamique dont lequel les auteurs jouent un rôle très important pour localiser les briques d'informations nécessaires.

9.3.3.3. Facilité d'utilisation

PrOALDoF-Web ne demande pas une forte connaissance dans le domaine informatique et technologique. Il propose une interface très simple. L'auteur de la formation a un espace de contrôle pour vérifier le contenu de la formation. Le reste est une boîte noire, il ne voit rien, car tout le traitement se fait à l'intérieur du système et il voit uniquement les résultats qui sont les sous fragments extraits à partir du Web qu'il doit les vérifier.

Un auteur avec les autres méthodes, est seul devant une diversité d'outils de recherche et multiple techniques et méthodologies de recherche. Un choix devant cette situation représente une tâche difficile à réaliser. Ensuite, l'auteur doit filtrer tous les résultats pour localiser ce qu'il a besoin. Les outils de création dans les plateformes ou bien dans les récentes recherche dans ce sens offrent des interfaces simples à utiliser, mais le problème majeur, c'est que les auteurs ne sont pas tous des experts dans la technologie et les 'TIC'.

9.3.3.4. Recherche du contenu

PrOALDoF-Web offre une méthode automatique pour la recherche des informations les plus pertinentes. Nous jugeons leurs pertinences en se basant sur une ontologie de domaine et en appliquant les mêmes filtres du moteur de recherche Google. Puis, nous comparons le contenu des informations en utilisant les règles sémantiques, ce qui augmente la pertinence des sous fragments Web par rapport aux sous parties du cours. Cela, implique une bonne qualité de recherche.

Dans les autres méthodes, si l'auteur est un expert il peut faire une bonne recherche rapide et efficace. Mais, si le contraire, la qualité de la recherche n'est pas assurée, surtout devant un immense nombre de pages Web existantes et devant un multiple choix de ressources existantes.

9.3.3.5. Expérience des auteurs

Si un auteur utilise le système PrOALDoF-Web, quelque soit son expérience et quelque soit son niveau, il n'est pas besoin d'avoir une forte expertise pour vérifier le contenu. Car, son rôle est le contrôle c'est tout, la qualité du contenu est assurée dès le début par les experts impliqués dans les différentes parties des approches proposées, dans la création des deux ontologies intégrées dans le système PrOALDoF-Web 'ontologie de domaine et ontologie pédagogique'.

Les autres systèmes exigent une certaine expérience des auteurs pour trouver les bonnes informations aux bons moments. Cela n'est pas assuré, car les auteurs n'ont pas ni les mêmes niveaux, ni les mêmes expertises et expériences. Pour cela, les cours créés ne sont pas toujours d'une meilleure qualité.

Les recherches actuelles dans le E-Learning proposent des approches basées sur l'intervention de certains experts, dans ce cas, le problème reste en position, car les auteurs vont créer eux-mêmes les cours via le système de l'approche proposée.

9.3.3.6. Qualité du contenu

Dans le cadre du projet PrOALDoF-Web, un ensemble d'experts ont participé dans la création des ontologies, une de domaine et une autre pédagogique, ainsi dans la définition des règles sémantiques. Un expert a de l'expérience pour juger la performance de l'approche proposée qui pose sur l'expérience qu'il a. Donc, la qualité des cours repose sur les différentes interventions des experts sur la manière de les créer.

Lors de la recherche dans notre projet, un premier filtrage des documents est assuré par la requête bien formulé par le système en se basant sur une ontologie de domaine. Puis, un autre filtrage du contenu sera effectué pour associer chaque fragment à une composante du cours. Enfin, un dernier filtrage concerne l'extraction des sous fragments par rapport aux règles sémantiques avec une association unique. Alors, notre système offre une forte qualité dans le sens où il trouve les bons sous fragments nécessaires dans la formation.

La qualité dans les autres méthodes de création n'est pas assurée à 100%. Par exemple, un auteur crée son cours et il le met sur une plateforme d'enseignement via les outils de création des cours. La qualité de ce cours n'est pas assurée car elle est liée à l'expérience de l'auteur. Il peut trouver des informations sur le Web, mais est-ce qu'elles sont fiables ou non c'est ça le problème. Donc, la qualité du contenu peut basculer entre l'intervalle 'Faible-Forte'.

9.3.3.7. L'alignement du but d'apprentissage

Dans PrOALDoF-Web, le but de la formation est fixé dans l'ontologie pédagogique qui présente un espace personnalisé de navigation pour un apprenant. Le contenu de la formation est défini aussi dans l'ontologie de domaine. Donc, la formation doit être adaptée au profil de l'apprenant tout au long son parcours dans le système.

Les autres méthodes ont un but bien précis comme dans les nouvelles approches d'adaptation. Mais pour les auteurs qui créent leurs cours selon un programme académique ou professionnel, ils n'arrivent pas à adapter les cours à tous les apprenants.

9.3.3.8. Feedback et adaptation

PrOALDoF-Web propose toute une approche d'adaptation du contenu en se basant sur l'ontologie pédagogique et sur les niveaux des apprenants. L'approche propose une adaptation personnalisée à chaque niveau abouti par un apprenant en sélectionnant les sous fragments nécessaires à partir de la base 'BNF'. En enchaînement, un feedback doit être effectué pour permettre au système de savoir ; à chaque test établi ; le niveau de l'apprenant. PrOALDoF-Web est un système adaptatif, il adapte les cours aux apprenants selon leurs retours d'expérience (feedbacks).

Dans la création manuelle, les auteurs créent les cours pour tous les apprenants, le contenu du cours est le même pendant toute la présentation et l'explication, cette dernière est adaptable en relation avec les auteurs et les apprenants. Le même cas dans les plateformes d'enseignement.

Pour le cas des recherches actuelles en adaptation, les chercheurs proposent des approches pour créer des systèmes éducatifs adaptatifs. Les approches contiennent des solutions aux feedbacks des apprenants dont le but, l'adaptation des cours.

9.3.3.9. Motivation

Dans la cadre du projet PrOALDoF-Web, nous avons créé une ontologie pédagogique qui contient pour chaque niveau un ensemble de règles sémantiques. Ces dernières sont divisées en deux catégories : pré-requis et informations supplémentaires. La seconde va motiver chaque apprenant à explorer les liens qui peuvent avoir d'autres explications et d'autres informations. En plus, l'adaptation est une approche basée sur l'expertise des experts, cela permet une bonne adaptation aux apprenants. Ces derniers vont avoir plus de motivation de continuer la formation en répondant à leurs besoins.

Dans les travaux de recherche actuels, les chercheurs veulent créer des systèmes assez performants pour pouvoir étudier les profils des apprenants en proposant des modèles apprenants avec plus de détails et plus de précision dans le sens avoir un bon suivi pour chaque apprenant.

Ils proposent aussi, des trucs motivants pour attirer l'attention des apprenants à utiliser un tel système. Le deuxième cas est absent dans les autres outils, car les auteurs n'ont pas un suivi et une adaptation permanente aux apprenants

9.3.3.10. Présentation du design

Elle n'est pas assurée à 100% dans PrOALDoF-Web car la présentation est uniquement textuelle. Notre proposition via ce projet se concentre sur le texte, ni visuel, ni audiovisuel sont entamés. La proposition est très difficile pour avoir un système de création automatique des cours textuellement. Donc, si nous améliorons le système par d'autres propositions pour autres formats d'informations, ça sera très intéressant pour un tel système.

Manuellement, la fonctionnalité est très faible dans le contenu des cours, sauf pour les auteurs qui maîtrisent bien les technologies récentes dans le visuel et audiovisuel.

Contrairement dans les recherches récentes et les plateformes puissantes où les professionnels améliorent d'un jour à l'autre les techniques d'amélioration du contenu vers d'autres formats qui facilitent la compréhension du système et augmente la satisfaction des apprenants via le design et les méthodes de présentation du contenu.

9.3.3.11. Ergonomie de l'interaction

Dans PrOALDoF-Web et quelques récents travaux, la navigation proposée par chacun de ces systèmes de formation est plus facile à utiliser avec une haute qualité de l'interface. Un aide est assuré pour les apprenants pour pouvoir documenter sur les plans de navigation et sur le principe du fonctionnement. Les systèmes peuvent être aussi utilisés dans des contextes d'apprentissage pour une diversité des apprenants, l'interface du cours se change en relation avec l'apprenant.

Par contre, les autres outils de création ne peuvent pas assurer cette fonctionnalité. Chez les auteurs en mode manuel, la fonctionnalité est nulle à cause de l'objectif technique du cours. Le cours est créé pour être présenté, donc le suivi et l'interaction avec le système ne sont pas assurés. Dans les plateformes, une navigation simple est assurée pour l'ensemble des cours disponibles. Mais, elles n'assurent pas l'enrichissement de l'interface de chaque cours par de nouveaux concepts liés au profil de l'apprenant, car elles ont une classification des cours uploadés par catégories.

9.3.3.12. Accessibilité

PrOALDoF-Web n'assure pas complètement l'accessibilité. C'est vrai que le système permet un contrôle dans l'interface par les concepteurs de la formation, ainsi, une présentation standard 'HTML' extensible à tout changement vers un format souhaité par un système d'exploitation, un appareil ou un navigateur. Mais, le problème dans notre approche, c'est que nous n'avons pas travaillé sur les systèmes d'enseignements pour les gens handicapés, car dans ce cas, il faut respecter quelques règles.

Dans certains travaux de recherche, les chercheurs se concentrent sur l'aspect accessibilité pour avoir des systèmes capables d'accepter un enseignement pour les handicapés, ou bien, pour d'accepter une version mobile du système afin d'améliorer les possibilités de connexion au système de formation. Manuellement (le mode manuel ou via les plateformes d'enseignement) les cours ne peuvent pas avoir une accessibilité, sauf s'ils sont intégrés dans les systèmes qui supportent cette fonctionnalité.

9.3.3.13. Réutilisation

PrOALDoF-Web et quelques récentes recherches, proposent des solutions de réutilisation du contenu de la formation pour avoir un apprentissage pour multiple contextes avec différents choix des apprenants. Dans PrOALDoF-Web la réutilisation est assurée par le billet de la base 'BNF' qui contient tout l'historique des sous fragments les plus pertinents qui jouent un rôle très important dans la création des cours adaptatifs aux apprenants.

Un exemple des travaux de recherche propose une ontologie contenant des briques d'informations de toute la formation pour un objectif la création dynamique du contenu du cours pour l'adapter au profil de l'apprenant.

Les autres outils, la création des cours s'effectue une seule fois lors l'objectif initial de la formation. La réutilisation du contenu n'est pas assurée car chaque cours se présente sur un seul bloc.

9.3.3.14. Conformité des normes standards du compliment

PrOALDoF-Web, les plateformes d'enseignement et quelques récentes recherches acceptent les normes internationales standards dans la création des cours. Dans un système E-Learning, il y a quatre niveaux des standards :

Le **premier** niveau vise à garantir l'ouverture et l'interopérabilité des systèmes 'OSI'. Le **deuxième** niveau se base sur les différents protocoles de communication 'HTTP' et d'ouverture des fichiers en ligne tels que : 'IETF' pour la messagerie, le 'FTP' pour le transfert des fichiers et le 'HTTP'. Le **troisième** niveau concerne l'encodage et l'indexation des données, les langages ('HTML', 'W3C' pour les formats 'HTML', 'XML', 'OWL', 'RDF', ...etc.). Enfin, le **quatrième** niveau porte sur la compatibilité entre les plates formes, les systèmes de gestion, les applications et les contenus d'E-Learning, pour l'ingénierie pédagogique et le suivi des apprenants. Dans ce dernier groupe, des standards spécifiques à l'enseignement en ligne ont été développés tels que : 'AICC', 'SCORM', 'IMS', 'DCMI', 'ARIADNE' et 'LOM'.

A titre d'exemple pour notre projet, l'utilisation d'une ontologie pédagogique est d'un format standard 'OWL'. Le cours se présente sous un format 'HTML' et le développement du système est réalisé par le langage de programmation 'JAVA', qui est reconnu par tous les systèmes d'exploitation. Pour le dernier standard utilisé en E-Learning : 'LOM', nous voulons dans les prochaines perspectives, intégrer notre projet dans la partie responsable à la création et à la gestion des cours : 'LMS' ou 'LCMS'.

Pour la création manuelle, un enseignant n'est pas obligé de suivre un tel niveau des standards. Dans ce cas, l'enseignant est libre dans ses choix et il crée ses propres cours selon le contexte et selon ses besoins ou bien selon l'autorité qu'il l'appartient. Sauf s'il utilise l'un des outils de création des cours dans les plateformes, car cet outil accepte les standards.

Pour mieux comprendre les résultats de la comparaison, nous réalisons les schémas des six cas des méthodes de création des cours avec les 14 points de comparaison (Voir Figure 22 en page 165).

9.4. Conclusion

PrOALDoF-Web présente un système de création des cours à partir du Web. Nous avons vu une partie du fonctionnement de la création des cours selon l'exemple illustratif présenté dans ce chapitre. L'exemple a montré deux phases très importantes dans PrOALDoF-Web : la fusion et l'adaptation. Les deux phases ont comme objectif la création des cours adaptatifs aux profils des apprenants. Une efficacité ce voit par cet exemple dans le sens que l'approche trouve les éléments nécessaires qui sont appropriés aux besoins de la formation et surtout que toute l'approche fonctionne d'une façon automatique.

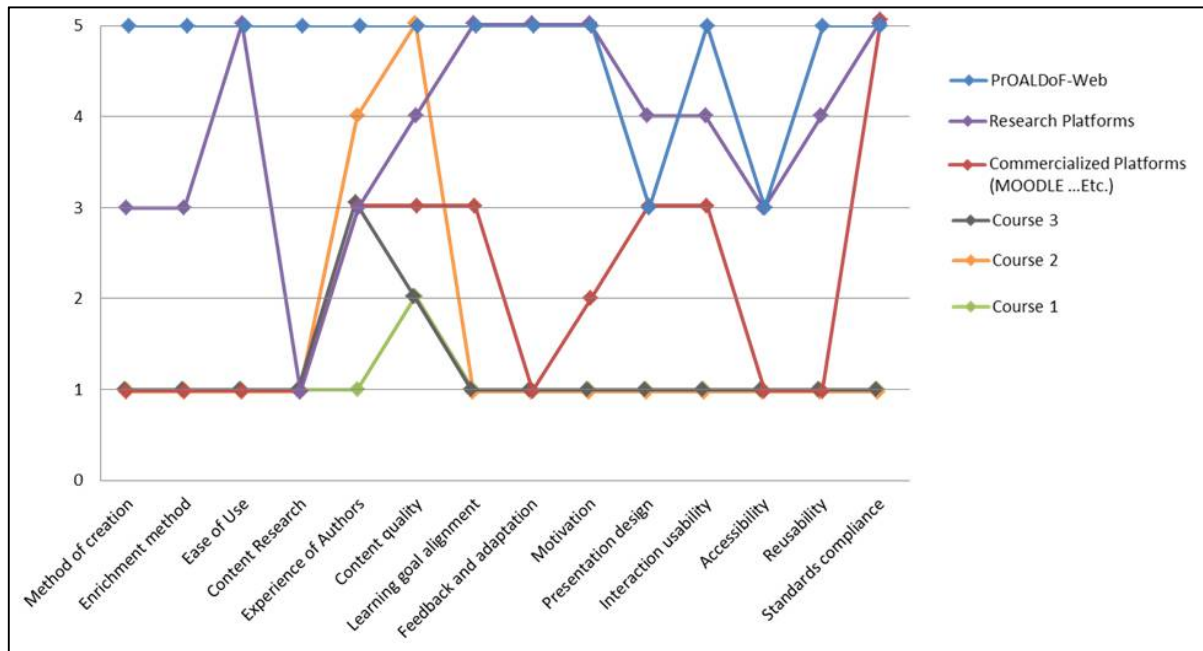


Figure 22 Diagramme des différentes méthodes de création des cours

Nous avons effectué l'analyse comparative avec les différentes méthodes de création des cours, pour montrer l'efficacité de l'approche proposée en se basant sur 14 points comparatifs. Depuis cette comparaison, nous avons constaté que notre système est capable de trouver les éléments indispensables dans la création des cours. Les éléments sont extraits à partir du Web.

Les deux ontologies créées dans le cadre du projet PrOALDoF-Web sont basées sur l'expertise humaine. Nous avons donc, un bon outil de création selon l'expertise fournie par des experts en domaine.

10. Conclusion générale

Le terme E-Learning est présent dans tous les secteurs de vie, il présente une forte demande dans les secteurs académiques, administratifs et industriels. E-Learning est basé sur un réseau informatique, que ce soit local ou mondial. Actuellement, ce type de formation représente un axe très important dans la recherche scientifique.

A travers PrOALDoF-Web, nous considérons qu'une formation en ligne 'E-Learning' est un système capable de trouver tous les éléments nécessaires à partir du Web pour créer automatiquement un contenu assez riche et d'une haute qualité. Ce contenu doit être sauvegardé pour pouvoir l'adapter aux profils et aux exigences des apprenants.

Cette nouvelle considération est due au manque d'un outil de création automatique des cours. Les outils existants obligent les auteurs et les concepteurs de la formation en ligne de consacrer un temps indéterminé et un effort très élevé surtout pour ceux qui ont moins d'expertise et moins d'expérience en recherche documentaire. Les auteurs équipés avec des moyens de recherche notamment sur le Web tels que les moteurs de recherche, les bibliothèques en ligne et les sites ou portails documentaires, sont devant un processus de recherche souvent très compliqué dont le but la découverte des ressources Web les plus pertinentes qui sont impliquées dans le processus de création des cours.

La très grande masse d'informations devenue disponible dans le Web, nécessite un bon filtrage et une fructueuse extraction. Dans un espace comme le Web, l'un des enjeux concernant l'éducation réside dans la création de contenus pédagogiques. Les enseignants sont confrontés à la nécessité de produire ou de numériser des matériaux éducatifs non seulement pour les présenter aux apprenants, mais aussi pour les rendre disponibles aux autres enseignants tout en utilisant des stratégies pertinentes pour les découvrir, les partager, les stocker et enfin les utiliser et réutiliser de multiples façons.

Parmi les outils utilisés souvent dans les universités, il y a les plateformes d'enseignement. Elles offrent des espaces pour mettre des contenus pédagogiques bien ordonnés pour être utilisés par les apprenants ou les enseignants. Ces plateformes intègrent des outils de création et de gestion du contenu de la formation tout en laissant la liberté du choix du contenu aux enseignants. Aucune plateforme ne peut créer un cours automatique, chaque cours est assuré par un enseignant responsable.

Les outils de création des cours fonctionnent de la même façon, ils proposent aux enseignants des fonctionnalités et des champs à remplir afin de faciliter la tâche de création. Mais, la recherche, le filtrage et l'extraction de l'information demandée doit être pris en charge par les auteurs de la formation.

PrOALDoF-Web a résolu ce problème major dans les étapes de création des cours, en proposant une approche basée sur des méthodes statistiques et sémantiques pour localiser l'information demandée. L'intégration d'une API Google permet l'utilisation permanente de la recherche automatique dans le moteur de recherche Google. Le processus de la recherche est basé essentiellement sur une formulation automatique des requêtes avec des bonnes combinaisons des mots-clés. Ces derniers sont proposés par des experts du domaine afin d'augmenter le taux des documents pertinents par rapport au domaine de recherche.

Une ontologie de domaine est créée au sein de ce projet. Elle contient tous les mots-clés indispensables dans toute formulation d'une requête de recherche. L'intégration de cette ontologie dans le système PrOALDoF-Web précède le processus de la recherche.

Un deuxième processus concerne le filtrage des documents jugés pertinents par rapport aux requêtes formulées. Ce processus représente une série de calculs, dont lesquels la découverte des morceaux nécessaires dans la création des cours joue un rôle très important. Le processus veut en premier lieu associer chaque document à une seule partie du cours présente dans l'ontologie de domaine en calculant la méthode du degré de pertinence 'DP'. Dans un deuxième lieu, le processus associe chaque sous fragment à une seule règle sémantique sauvegardée dans l'ontologie de domaine en se basant sur la méthode 'DBRS'. Une règle sémantique est une suite de termes techniques dans une notion d'une partie du cours. L'ensemble des règles sémantiques composent une partie du cours.

Cette méthode est composée de trois sous méthodes 'Formules'. Nous calculons d'abord la formule 'DBRS_M' qui présente la moyenne des termes existants dans un sous fragment par rapport aux termes d'une règle sémantique divisé par le nombre total des mots du sous fragment. Puis, la formule 'DBRS_D' qui présente la distance entre les termes existants. Enfin, la formule 'DBRS_{PA}' qui présente une précision de l'association d'un sous fragment à une règle sémantique où chaque association est unique. Cette précision est basée sur les deux formules précédentes, elle égale à la multiplication de la moyenne 'DBRS_M' avec le reste de la division du nombre total des termes existants par la distance 'DBRS_D'.

Nous avons réussi à travers notre projet PrOALDoF-Web donc, à proposer une approche de création d'un domaine d'apprentissage sans aucune intervention des concepteurs 'enseignants'. L'approche commence par : trouver les documents les plus pertinents qui se trouvent sur le Web, les téléchargés, les filtrés en se basant sur notre ontologie de domaine et sur nos règles sémantiques, enfin extraire les parties les plus pertinentes en appliquant la méthode de fusion proposée afin de créer le domaine d'apprentissage de l'approche proposée. Par la suite, une adaptation basée fusion sera effectuée pour créer les cours dynamiquement à partir du domaine d'apprentissage créé, dont le but, obtenir un cours adaptatif au profil de l'apprenant.

Donc, notre approche proposée permet de réduire l'immense espace du Web qui contient des milliards de pages Web. Cet espace devient personnalisé avec l'adaptation directe des cours aux profils des apprenants afin de répondre à leurs besoins et afin d'augmenter leurs satisfaction en fournissant une bonne formation extensible à tout changement et mise à jour, mais avec des ressources éducatives fiables.

La fusion dans PrOALDoF permet de construire un cours à partir d'un ou plusieurs documents Web. Le but est d'augmenter au maximum la qualité des cours créés dans le système. Cette fusion représente un processus de sélection à partir des sous fragments les plus pertinents sauvegardés dans notre base 'BNF', les sous fragments ayant la précision 'DBRS_{PA}' la plus élevée. Chaque ligne d'une règle sémantique dans la base 'BNF' est composée de plusieurs sous fragments ordonnées selon leurs 'DBRS_{PA}' par rapport à cette règle.

Le processus d'adaptation dans PrOALDoF-Web permet la création d'un domaine d'apprentissage adaptatif. Il est basé sur le principe de la fusion en sélectionnant les sous fragments appropriés aux règles sémantiques définies pour chaque niveau qui peut être atteint par les apprenants. Les niveaux sont sauvegardés dans l'ontologie pédagogique créée dans le cadre de notre projet.

Donc, nous pouvons confirmé que PrOALDoF-Web est un système d'enseignement en ligne dont le contenu pédagogique est adaptatif au profil de l'apprenant et dont la source d'information disponible sur le Web est bien entendu extraite et réorganisée selon un canevas prédéfini du domaine d'enseignement 'Modèle Excel' afin d'enrichir convenablement à partir de la base 'BNF' les connaissances de l'apprenant selon son niveau.

“**Proposition and Organization of an Adaptive Learning Domain based on Fusion from the Web**” : est l’abréviation technique de PrOALDoF-Web. Ce système a répondu aux objectifs cités dans l’introduction générale. Il est composé de trois parties : Partie Recherche, Partie Fusion et Partie Adaptation. Les trois parties peuvent travailler en simultané ou individuellement, mais à la fin du processus de traitement les trois parties devenues complémentaires pour créer les cours adaptatifs.

Nous avons vu une partie du fonctionnement de la création des cours dans l’exemple illustratif présenté dans le chapitre d’évaluation du système. Nous avons montré l’efficacité de notre système via l’exemple en trouvant les sous fragments les plus pertinents pour quelques règles sémantiques.

Les résultats de l’évaluation reflètent le succès du système PrOALDoF-Web en parlant de la performance trouvée selon la comparaison effectuée avec les différentes méthodes de création des cours qui sont mentionnées dans notre travail. Cette comparaison est basée sur 14 points comparatifs. Depuis cette comparaison, nous avons constaté que notre système est capable de trouver les éléments indispensables dans la création des cours. Les éléments sont extraits à partir du Web.

Les deux ontologies créées dans le cadre du projet PrOALDoF-Web sont basées sur l’expertise humaine. Nous avons donc, un outil complet de très bonne qualité de création selon l’expertise fournie par des experts en domaine et selon toutes les méthodes proposées dans ce cadre.

PrOALDoF-Web offre de nouvelles techniques qui ont transformées le Web comme étant un support documentaire des ressources qui sont bien organisées et adaptées directement aux besoins des apprenants. L’adaptation des ressources Web extraites, donne la fiabilité, la satisfaction des apprenants et la qualité du contenu dans la plateforme E-Learning. Ce système n’a pas besoin de beaucoup d’efforts fournis et d’un temps pris qui peut être indéterminé, les deux concernent les auteurs. Le système crée des cours qui peuvent obtenir une grande qualité parce que leurs contenus sont obtenus à partir de la grande richesse du Web. La mise à jour des cours se fait directement à partir du Web avec la réutilisabilité des ressources extraites sauvegardées dans la base ‘BNF’.

11. Perspectives

PrOALDoF-Web demande une forte expertise des experts impliqués dans la création des règles sémantiques. L'expertise est surtout dans le côté technique et linguistique. Le but est pour trouver toutes les informations appropriées par rapport aux règles sémantiques. Donc, ces dernières doivent être bien formulées et structurées afin d'obtenir les meilleurs résultats avec une efficacité améliorée du système. Mais, le nombre des règles sémantiques est très important pour chaque composante du cours. Pour cela, nous pensons dans nos futures recherches à travailler dans ce sens en proposant une méthode automatique de création des règles sémantiques.

Nous avons traité à travers PrOALDoF-Web que des formats textuels. Donc, un deuxième besoin concerne le traitement de d'autres formats d'informations (Image, Multimédia ...etc.).

Enfin, un troisième besoin concerne la standardisation de l'approche proposée en l'intégrant dans les standards des systèmes d'enseignement en ligne.

PARTIE - III : ANNEXES

Contenu

L'environnement et les outils utilisés

12. L'environnement et les outils utilisés

12.1.Introduction

Nous présentons dans cette partie l'environnement de programmation utilisée, ainsi, tous les outils et toutes les API utilisées dans le cadre du projet PrOALDoF-Web.

12.2.L'environnement Java : NetBeans

12.2.1. Présentation

NetBeans est un projet open source fondé par Sun Microsystems. L'IDE NetBeans est un environnement de développement permettant d'écrire, compiler, déboguer et déployer des programmes. Il est écrit en Java – mais peut supporter n'importe quel langage de programmation. Il y a également un grand nombre de modules pour étendre l'IDE NetBeans (Meniet, 2007).

Il est développé par Sun en juin 2000 sous licence CDDL et GPL v2 (Common Development and Distribution License et GENERAL PUBLIC LICENSE Version 2). En plus de Java, NetBeans permet également de développer avec d'autres langages tels que : Python, C, C++, Ruby, XML, PHP et HTML (Ubuntu, 2012).

Il comprend toutes les caractéristiques d'un IDE moderne (coloration syntaxique, projets multi-langage, refactoring, éditeur graphique d'interfaces et de pages web, ...etc.) [(Ubuntu, 2012) et (NetBeans, 2013)].

Conçu en Java, NetBeans est disponible sous Windows, Linux, Solaris et plusieurs d'autres systèmes d'exploitation (NetBeans, 2013).

12.2.2. Historique

En 1997, NetBeans naît de Xelfi, un projet d'étudiants dirigé par la Faculté de mathématiques et de physique de l'Université Charles de Prague. Plus tard, une société se forme autour du projet et édite des versions commerciales de l'EDI NetBeans, jusqu'à ce qu'il soit acheté par Sun en 1999. Sun place le projet sous double licence CDDL et GPL v2 en juin de l'année suivante (NetBeans, 2013).

12.2.3. Installation

L'IDE NetBeans est un produit gratuit, sans aucune restriction quant à son usage. Vous pouvez télécharger l'IDE NetBeans à partir du site officiel : <http://www.netbeans.org/>.

Figure 23 ci-dessous présente la page principale de NetBeans.

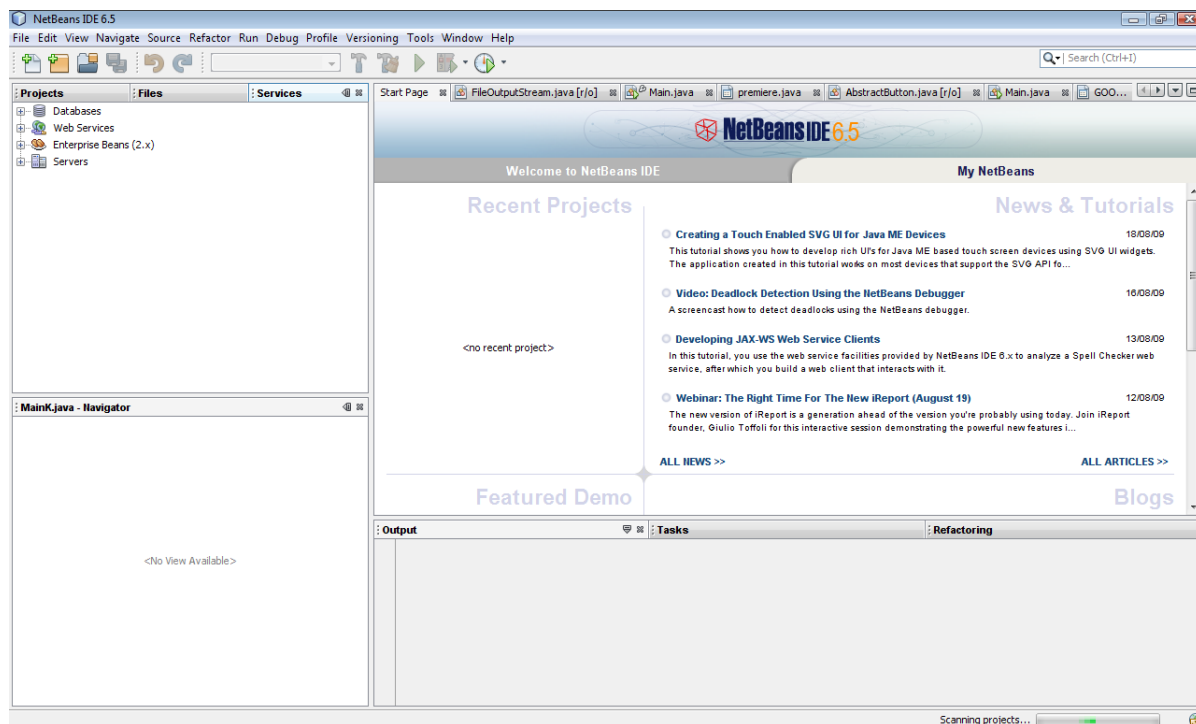


Figure 23 La page principale de NetBeans IDE 6.5

12.3. Protégé2000

12.3.1. Présentation

Protégé est le plus connu et le plus utilisé des éditeurs d'ontologie. Open-source, développé par l'Université de Stanford, il a évolué depuis ses premières versions (Protégé-2000) pour intégrer à partir de 2003 les standards du Web sémantique et notamment OWL. Il offre de nombreux composants optionnels : raisonneurs, interfaces graphiques ...etc. (PUS, 2013c).

Protégé est un éditeur qui permet de construire une ontologie pour un domaine donné, de définir des formulaires d'entrée de données, et d'acquérir des données à l'aide de ces formulaires sous forme d'instances de cette ontologie. Protégé offre un environnement graphique interactif pour la conception d'ontologies. Un arbre permet une navigation rapide et simple dans la hiérarchie (PUS, 2013b).

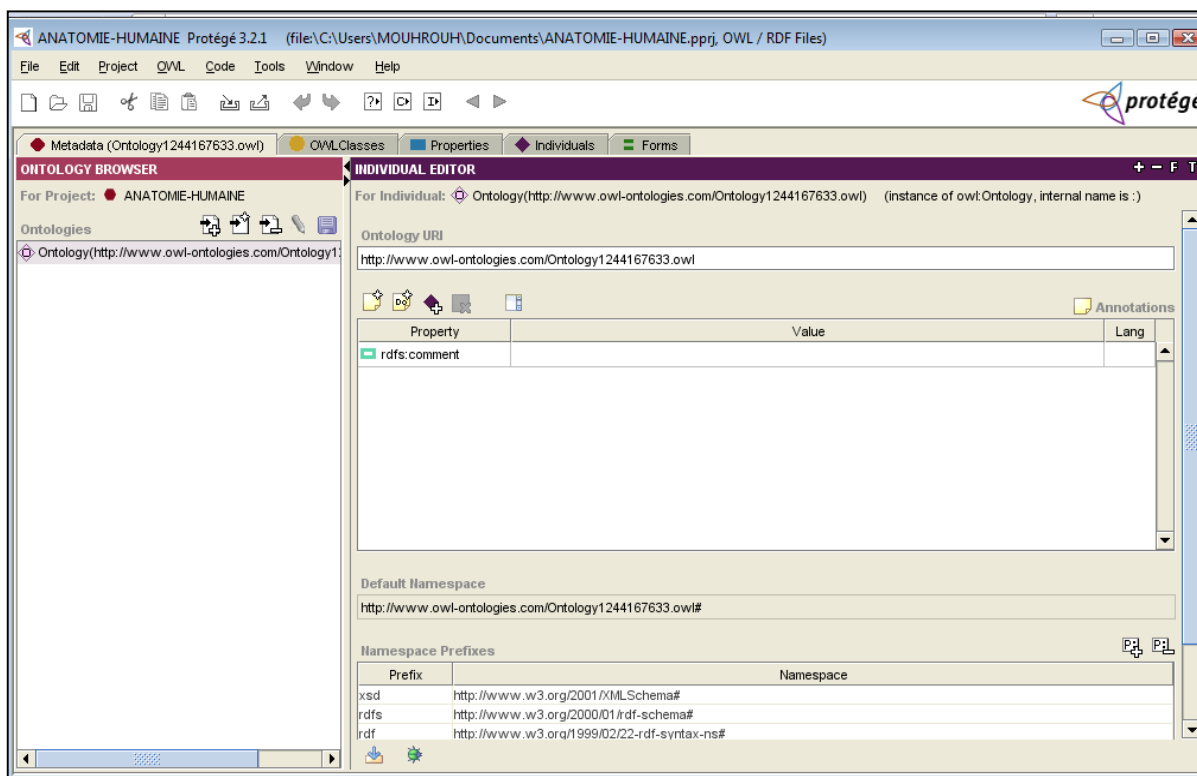


Figure 24 Projet ouvert sous Protégé v 3.2.1

Le modèle de Protégé est basé sur les frames. Il aide à définir les classes et les hiérarchies avec héritage multiple ; les attributs, les restrictions de valeurs de ces attributs, leurs facettes, comme les restrictions de cardinalité, valeurs par défaut, ainsi que des attributs inverses, des méta-classes et la hiérarchie de méta-classes.

Il faut noter que des interfaces à l'UMLS™ [Système médical unifié de langue, en anglais : Unified Medical Language System,] (Lindberg et al, 1993) et Word Net permettent aux utilisateurs d'intégrer et importer des éléments de ces grandes sources de connaissances online dans leurs ontologies (PUS, 2013b).

Protégé est également une librairie Java qui peut être étendue pour créer de véritables applications à bases de connaissances en utilisant un moteur d'inférence pour raisonner et déduire de nouveaux faits par application de règles d'inférence aux instances de l'ontologie et à l'ontologie elle-même (Méta-raisonnement). Nous pouvons dire qu'une ontologie est une représentation des connaissances d'un domaine (PUS, 2013b).

La Figure 24 ci-dessus présente un projet ouvert sous la version Protégé 3.2.1, ainsi quelques composantes de cet outil.

Protégé2000 est une interface modulaire permettant l'édition, la visualisation, le contrôle (vérification des contraintes) d'ontologies, l'extraction d'ontologies à partir de sources textuelles, et la fusion semi-automatique d'ontologies. Le modèle de connaissances sous-jacent à Protégé2000 est issu du modèle des frames et contient des classes (concepts), des slots (propriétés) et des facettes (valeurs des propriétés et contraintes), ainsi que des instances des classes et des propriétés (PUS, 2013c).

Protégé2000 se veut compatible avec le modèle OKBC 'Ouvrir la connectivité base de connaissance, en anglais : Open Knowledge Base Connectivity', mais en diffère sur certains points. Par exemple, un frame, qui est soit une classe, soit un slot, soit une facette, est instance d'une classe et une seule dans Protégé2000 ce qui n'est pas toujours le cas dans OKBC, où un même frame peut d'ailleurs représenter plusieurs objets conceptuels. Protégé2000 autorise la définition de méta-classes, dont les instances sont des classes, ce qui permet de créer son propre modèle de connaissances avant de bâtir une ontologie (PUS, 2013c).

12.3.2. Les différentes fonctionnalités de Protégé 2000

- Construit interactivement et importe des ontologies.
- Une bibliothèque d'ontologies est disponible sur le site officiel (PUS, 2013b).
- Programmer et incorporer des plugins.
- Fonction d'export vers différent format de description d'ontologie : 'RDF(S), OIL, DAML+OIL, OWL' via plug-in pour certains formats.

12.3.3. Installation

On peut installer facilement Protégé 2000 dans les différents systèmes d'exploitation tel que (Windows, Linux, Solaris ...etc.). Mais il faut d'abord télécharger une version de l'outil gratuitement via le site officiel de Protégé (PUS, 2013a).

12.4. Api Google

12.4.1. Définition

L'API Google est un kit de développement logiciel disponible librement, qui permet de créer de nouvelles applications utilisant directement la base de données des pages indexées par Google, par le biais d'un service web (Duffez, 2002).

Une API a pour objet de faciliter le travail d'un programmeur en lui fournissant les outils de base nécessaires à tout travail à l'aide d'un langage donné. Elle constitue une interface servant de fondement à un travail de programmation plus poussé (Roussel, 2009).

12.4.2. Description

Une API pour la programmation dans un langage x est ainsi composée d'un ensemble de fonctions, routines et méthodes, écrites dans ce langage. Ces fonctions de bas niveau (ouvrir un fichier, le parser, le fermer, ...) ont la propriété d'être d'un usage courant dans toutes les applications dérivées de x. D'où l'idée de les programmer une fois pour toutes puis de les mettre à disposition de la communauté des programmeurs. Ces derniers n'ont plus alors qu'à choisir et appeler les fonctions pertinentes selon les objectifs de leur code.

Citons par exemple les API SAX (Simple API for XML) et DOM (Document Object Model) pour la programmation XML, ou les API Java pour la programmation Java RMI, Java IDL, Java Media API, ...etc.).

12.4.3. Les étapes pour utiliser l'API Google

C'est vraiment très simple, il vous suffit de connaître un peu le langage ou vous voulez intégrer l'API et de suivre les 3 étapes suivantes (Duffez, 2002) :

- **Étape 1 : s'inscrire auprès de Google et obtenir une clé**

Pour utiliser l'API Google, il faut s'enregistrer auprès de Google. Cette procédure est très simple, il suffit de donner un mot de passe et une adresse email (il est obligatoire que la boîte soit Gmail) dans le formulaire situé sur : <https://www.google.com/accounts/Login>. Après confirmation, vous recevrez votre clé ('key') qui servira à développer avec l'API Google.

- **Étape 2 : récupérer les classes d'accès au langage**

Pour développer des outils, vous avez besoin d'une série de classes écrites au langage de l'API téléchargée, qui permettent d'interroger très facilement Google. Vous pouvez télécharger l'API dans le langage souhaité, par exemple si nous voulons travailler avec Java, il faut télécharger l'API Google Java.

- **Etape 3 : développer ses propres outils**

Une fois que vous avez obtenu votre clé et récupéré les classes, il est très simple de faire une requête Google. Par exemple, pour intégrer la clé obtenue il faut écrire : Google key = '01234567890123456789012345678901'; l'instruction dépende au langage sélectionné.

Libre à vous de faire d'autres requêtes, ça dépend de votre besoin, Sachez cependant que Google limite l'utilisation de la Google API à 1000 requêtes par jour (chaque requête pouvant renvoyer au maximum 10 résultats).

12.4.4. Fonctionnement

Les utilisateurs écrivent des programmes qui seront connectés d'une façon ou d'une autre à Google Web API service. Ce dernier va effectuer une communication avec le serveur de Google par l'intermédiaire d'un protocole SOAP 'Simple Object Access Protocol', pour le traitement des informations entrées par le programme utilisateur (El Achbor et El Korchi, 2005). Le moyen de communication avec Google Web APIs service peut être un programme CGI 'Common Gateway Interface'.

12.4.5. Liste des API Google

On présente quelques API Google, les plus utilisées (Duffez, 2006) :

- API Blogger
- API Google AdWords
- API Google AdWords Sandbox
- API Google Calendar
- API Google Checkout
- API Google Deskbar (ou Google Desktop)
- API Google Gadgets
- API Google Homepage
- API Google Maps
- API Google Search
- API Google Suggest
- API Google Talk
- API Google Toolbar
- Google AJAX Search API

Celle que nous avons utilisé, c'est l'API Google Search, avec le langage Java (environnement NetBeans).

12.5.Api Excel

12.5.1. Description

Java Excel API est un kit de développement logiciel, Open Source, permettant de lire, d'écrire, et de modifier dynamiquement les feuilles de calcul Excel. Maintenant les développeurs Java peuvent lire des feuilles de calcul Excel, de les modifier avec une API simple et commode, et écrire les changements à n'importe quel flux de sortie (par exemple, disque, HTTP, base de données, ou toute autre prise de courant) (JEA, 2011b).

Tout système d'exploitation qui peut exécuter une machine virtuelle Java (à savoir, pas seulement Windows) peut à la fois développer et atteindre les feuilles de calcul Excel. Parce qu'il est de Java, l'API peut être appelée dans une servlet, donnant ainsi accès à des feuilles de calcul Excel, grâce à internet et intranet 'applications web' (JEA, 2011b).

L'API est disponible sur le site officiel de SourceForge 'sourceforge.net' (JEA, 2011a).

12.5.2. Quelques caractéristiques

On peut citer quelques caractéristiques de cette API (JEA, 2011b) :

- Lire les données des classeurs Excel 95, 97, 2000, XP et 2003 classeurs ;
- Lire et écrire les formules (Excel 97 et versions ultérieures uniquement) ;
- Générer des feuilles de calcul en format Excel 2000 ;
- Prise en charge de la police, numéro et date de mise en forme ;
- Supporte l'ombrage, à la frontière, et la coloration des cellules ;
- Modifie les feuilles de calcul existantes ;
- Elle est internationalisé, le traitement favorable dans presque tous les paramètres régionaux, pays, langue, ou de codage de caractères (les formules sont actuellement pris en charge uniquement en anglais, français, espagnol et allemand, mais l'API permet d'ajouter d'autres langues si elles sont traduites) ;
- Prend en charge la copie de tableaux ;
- Prend en charge l'insertion et la copie des images dans les feuilles de calcul ;

12.6. Api JENA

12.6.1. Description

JENA est un Framework Java pour la construction sémantique des applications Web. Il fournit un environnement de programmation pour RDF, RDFS et OWL, SPARQL (langage de requêtes pour RDF) et comprend une règle basée sur un moteur d'inférence (JENA, 2011a).

JENA est open source, et développée à partir des travaux réalisés dans le laboratoire centrale de recherche de HP 'Hewlett-Packard' (JENA, 2011a), dans le domaine 'Recherche sur le Web sémantique' ou 'Semantic Web Research' (HP, 2011).

L'API JENA est disponible sur le site officiel de SourceForge 'sourceforge.net' (JENA, 2011b).

12.6.2. Les composantes de JENA

Le Cadre de JENA comprend (JENA, 2011a) :

- Une API RDF ;
- Lecture et écriture de RDF en RDF / XML, N3 et N-Triples ;
- Une API OWL ;
- Une mémoire persistante et un stockage ;
- Un moteur de requête SPARQL.

Références

- AD-APTE.** (2013). *AD-APTE Formation*. En ligne : <http://adapte-formation.fr/nous-connaître/ad-apte-formation/>
- AEC** : Aquitaine Europe Communication. (2005). *Dossier de veille AEC N°13 Moteurs de recherche*.
- AERIS** : Aide aux Étudiants pour la Recherche d'Information Scientifique. (2006). *Glossaire sur la recherche d'information sur Internet*. En ligne : http://aeris.11vm-serv.net/cours/util_outils/glos.html
- Akoumianakis, D.** (2011). *Learning as 'Knowing': Towards Retaining and Visualizing Use in Virtual Settings*. *Educational Technology & Society*, 14(3), 55–68.
- Allan (editor), J., Aslam, J., Belkin, N., Buckley, C., Callan, J., Croft (editor), B., Dumais, S., Fuhr, N., Harman, D., Harper, D-J., Hiemstra, D., Hofmann, T., Hovy, E., Kraaij, W., Lafferty, J., Lavrenko, V., Lewis, D., Liddy, L., Manmatha, R., McCallum, A., Ponte, J., Prager, J., Radev, D., Resnik, P., Robertson, S., Rosenfeld, R., Roukos, S., Sanderson, M., Schwartz, R., Singhal, A., Smeaton, A., Turtle, H., Voorhees, E., Weischedel, R., Xu, J., & Zhai, C-X.** (2002). *Challenges in Information Retrieval and Language Modeling . Report of a Workshop held at the Center for Intelligent Information Retrieval*, University of Massachusetts Amherst, 1-17.
- Amaral, L., & Leal, D.** (2006). *From classroom teaching to e-learning: the way for a strong definition*. In proceedings of WMSCI - World Multi-Conference on Systemics, Cybernetics and Informatics,(Ed.), DSI - Sociedade da Informação. Orlando, USA.
- Arthurs, J.** (2007). *A Juggling Act in the Classroom: Managing Different Learning Styles*. *Teaching and Learning in Nursing*, 2, (1), 2–7.
- Bachimont, B.** (2000). *Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances*. In Z. M. Charlet J., Kassel G., Bourgault D., (Ed.), *Ingénierie des connaissances. Évolution Récentes et nouveaux défis* Paris: Eyrolles, 305-323.
- Barbe, W-B., Swassing, R-H., & Milone, M-N.** (1979). *Teaching through modality strengths : concepts and practices*, Columbus, Ohio: Zaner-Bloser.
- Barbe, W-B., Swassing, R-H., & Milone, M-N.** (1988). *Teaching through modality strengths : concepts and practices*, (2^e ed.). Columbus, Ohio: Zaner-Bloser.
- Bastien, J.M.C. & Scapin, D.L.** (1993). *Critères Ergonomiques pour l'Évaluation d'Interfaces Utilisateurs (version 2.1)*. Technical report Ndeg.156, May 1993. INRIA. Programme 3 Artificial intelligence, cognitive systems, and man-machine interaction.
- Bastien, J. M. C., Leulier, C., et Scapin, D.** (1998). *L'ergonomie des sites web*. In Le Moal, J-C., et Hidoine, B.. Chapitre rédigé dans le cadre du projet européen Commerce & Interactions (EP 22287). INRIA.

- Bastien, J.M.C., Leulier, M. et Scapin, D.L.** (1998). *L'ergonomie des sites Web*. in *Créer et maintenir un service Web*, cours INRIA 28 Sept — 2 Oct 1998, Pau (France), ADBS Editions : Paris.
- Benadi, S. A.** (2004). *Structuration des données et des services pour le télé-enseignement*. Thèse de Doctorat, Présentée devant L'Institut National des Sciences Appliquées de Lyon, France.
- Benayache, A.** (2005). *Construction d'une Mémoire Organisationnelle de formation et évaluation d ans un contexte e-learning: le projet MEMORAe*, these en France.
- Berners-Lee, T.** (2000). *What the semantic web can represent*. En ligne : <http://www.w3.org/DesignIssues/RDFnot.html>
- Berners-Lee, T., Hendler, J., & Lassila, O.** (2001). *The Semantic Web. Scientific American*, 284(5), 34-43.
- Bibeau, R.** (1998). *Quand Dewey 'surfe' sur le web*. La bibliothèque scolaire à l'heure d'Internet.
- Bonnet, J.** (2013). *Les dix chiffres clés du net en 2012*. En ligne : <http://www.usinenouvelle.com/article/les-dix-chiffres-cles-du-net-en-2012.N190409>
- Borst, W. N.** (1997). *Construction of Engineering Ontologies*. Center for Telematica and Information Technology, University of Twente, Enschede, NL.
- Bostrom, R. P., Olfman, L., & Sein, M. K.** (1990). *The importance of learning style in end-user training*. MIS Quarterly, 14(2), 101–119.
- Boutemedjet, S.** (2004). *Web Sémantique et e-Learning 'Cours IFT6261'*, Sherbrooke – Canada.
- Brandon-Hall.** (2013). Le site officiel. En ligne : <http://www.brandon-hall.com/>
- Brooks, C., Kettel, L., & Hansen, C.** (2005). *Building a learning object content management system*. In G. Richards (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005*, (pp. 2836-2843). Chesapeake, VA: AACE. 24–28 October, Vancouver, Canada.
- Brusilovsky, P.** (1996). *Methods and techniques of adaptive hypermedia, Adaptive Hypertext and user Modeling and user adapted interaction*, V6 n° 2-3, 87-129.
- Brusilovsky, P.** (1999). *Adaptive and Intelligent Technologies for Web-based Education*. In C. Rollinger and C. Peylo (Eds.), *Special Issue on Intelligent Systems and Teleteaching*, 4, 19–25. *Künstliche Intelligenz*.
- Brusilovsky, P.** (2001). *Adaptive Hypermedia. Adaptive Hypertext and user Modeling and user adapted interaction*. V11, 87-110.
- Brusilovsky, P.** (2004). *Knowledge Tree: A distributed architecture for adaptive e-learning*. In *proceedings of WWW 2004*, (pp. 104–113), 17–22 May, New York, USA.

- BULCO** : Bibliothèque de l'Université du Littoral Côte d'Opale. (2011). *Méthodologie documentaire, La Recherche documentaire sur internet*, 1-18.
- BULCO** : Bibliothèque de l'Université du Littoral Côte d'Opale. (2012). *Méthodologie documentaire, LA RECHERCHE DOCUMENTAIRE SUR INTERNET*, 1-22.
- Canales, A., Pena, A., Peredo, R., Sossa, H., & Gutierrez, A.** (2007). *Adaptive and intelligent Web based education system: Towards and integral architecture and framework*. *Expert Systems with Applications*, 33(4), 1076–1089.
- Caravantes, A., & Galán, R.** (2011). *Generic Educational Knowledge Representation for Adaptive and Cognitive Systems*. *Educational Technology & Society*, 14(3), 252–266.
- Chabert-Ranwez, S.** (2000). *Composition automatique de documents hypermédias adaptatifs à partir d'ontologies et de requêtes intentionnelles de l'utilisateur*. Thèse de Doctorat, Université de Montpellier II, France.
- Chaoui, M., & Laskri, M.T.** (2011a). *New method of finding information on the Web in unstructured information resources for educational use by learners*. *International Journal of Research and Reviews in Computer Science (IJRRCS)*, 2(1), 33–39, Science Academy Publisher, United Kingdom.
- Chaoui, M., & Laskri, M-T.** (2011b). *Towards the Creation of Adaptive Content from Web Resources in an E-Learning Platform to Learners Profiles*. *International Journal of World Academy of Science, Engineering and Technology WASET*, 77 (27), 157–162.
- Chaoui, M., & Laskri, M. T.** (2013). *Proposition and Organization of an Adaptive Learning Domain based on Fusion from the Web*. *Educational Technology & Society*, 16 (1), 118–132.
- Charlier, B., & Peraya, D.** (2003). *Technologie et innovation en pédagogie : dispositifs innovants de formation pour l'enseignement supérieur*. De Boeck, 230 p, Bruxelles - Belgique.
- Chen, C.** (2008). *Intelligent Web-based learning system with personalized learning path guidance*. *Computers & Education*, 51(2), 787–814.
- Chen, C.** (2009). *Personalized E-Learning System with Self-Regulated Learning Assisted Mechanisms for Promoting Learning Performance*. *Expert Systems with Applications*, 36, 8816–8829.
- Chi, Y.** (2009). *Ontology-based curriculum content sequencing system with semantic rules*. *Expert Systems with Applications*, 36(4), 7838–7847.
- Chikh, A.** (2004). *Une méthodologie de réutilisation en ingénierie du document : Le système SABRA*. Thèse de Doctorat d'état en Informatique, INI, Alger, Algérie.
- Kyung Won Chung, K-W.** (1996). *Anatomie humaine*, Livre.
- Delestre, N.** (2000). *METADYNE, un Hypermédia Adaptatif Dynamique pour l'Enseignement*. Thèse de l'Université de Rouen, France.

- Dessonnaz, J-D.** (2013) *Filtrage de contenu*. FRI-TIC. En ligne : <http://www.fri-tic.ch/dyn/9.asp?url=41278.asp>
- Dico du Net** : est un dictionnaire collaboratif en ligne. (2013). En ligne : <http://www.dicodunet.com>
- Dillenbourg, P. & Martin-Michiellot.** (1994). *Le rôle des techniques d'intelligence artificielle dans les logiciels de formation*, Learntec, Karlsruhe.
- Jérôme DINET, J., ROUET, J-F. & PASSERAULT, J-M.** (1998). *LES ' NOUVEAUX OUTILS ' DE RECHERCHE DOCUMENTAIRE SONT-ILS COMPATIBLES AVEC LES STRATÉGIES COGNITIVES DES ÉLÈVES ?* Quatrième colloque hypermédias et apprentissages, 15, 16, 17 Octobre 1998 – POITIERS, 149-162.
- Duffez, O.** (2002). *Développer un outil avec la Google API*. En ligne : <http://www.webrankinfo.com/dossiers/api-google/api-soap>
- Duffez, O.** (2006). *Liste de toutes les API de Google*. En ligne : <http://www.webrankinfo.com/actualites/200601-liste-api-google.htm>
- Duffez, O.** (2008). *Google a répertorié 1000 milliards de pages sur le web !*, <http://www.webrankinfo.com/actualites/200807-1000-milliards-de-pages-sur-le-web.htm>
- Edmonds, E.A.** (1981). *Adaptive Man-Computer Interfaces*. In : Computing skills in the user interface, Coobs M.J & Alty J.L Eds. Computer and People series, New york: Academic Press, 1981, pp 389-426.
- edutechwiki.** (2006). *Adaptation*. En ligne : <http://edutechwiki.unige.ch/fr/Adaptation>
- El Achbor, B., & El Korchi, T.** (2005). *L'API DE GOOGLE*, Rapport de mini-projet. Caen - France.
- Le glossaire des ENMI12** : *Les Entretiens du nouveau monde industriel 2012*. En ligne : <http://www.enmi12.org/glossaire/>
- Farquhar, A., Fikes, R., & Rice, J.** (1996). *The Ontolingua Server: Tool for Collaborative Ontology Construction*. Paper presented at the Proc. of the 10th Knowledge Acquisition for Knowledge-Based Systems Workshop, Alberta, Canada.
- Fenineche, K.** (2007). *Adaptation des situations d'apprentissage dans les communautés de pratique des enseignants Approche basée sur les standards du e-learning*. Thèse de Magister, Alger, Algérie.
- Fernandez-Lopez, M., Gómez-Perez, A., & Juristo, N.** (1997). *Methontology: From Ontological Art Toward Ontological Engineering*. Paper presented at the Spring Symposium Series on Ontological Engineering. AAAI97, Stanford, USA.
- Gandon, F.** (2006). *Ontologies informatiques*, image, disponible sur : http://interstices.info/jcms/c_17672/ontologies-informatiques?part=3

- García-Barrios, V.M.** (2006). *Adaptive E-Learning Systems: Retrospection, Opportunities and Challenges*. In Proceedings of the 28th International Conference on Information Technologies Interfaces ITI 2006,(pp. 53–58), June 19-22,Cavtat, Croatia.
- Ouerfelli, T., & Gharbi, K.** (2008). *LE DISPOSITIF D'ENSEIGNEMENT A DISTANCE A L'UNIVERSITE DE BAHREIN , PRATIQUES ET ATTENTES DES ENSEIGNANTS*. En ligne : <http://docs.google.com/isdm.univ-tln.fr/PDF/isdm32/isdm32-gharbi.pdf>
- Gómez-Pérez, A.** (1999a). *Ontological Engineering: A state of the art*. Expert Update, 2(3), 33-43.
- Gómez-Pérez, A.** (1999b). *Tutorial on Ontological Engineering*. Paper presented at the Proc. IJCAI99.
- Grari, M.** (2007). *VEILLE TECHNOLOGIQUE SUR INTERNET*. 1-52. En ligne : <http://esto.ump.ma/mounir/cours/veille/veille%20Tech.pdf>
- Guarino, N., & Welty, C.** (2000). *A formal Ontology of Properties*. Proceedings of the *ECAI-00 Workshop on Applications of Ontologies and Problem-Solving Methods*, Berlin, pp. 12/1-12/8.
- Guarino, N., & Welty, C.** (2000). *Identity, Unity and Individuality: Towards a Formal Toolkit for Ontological Analysis*. Proceedings of the *14th European Conference on Artificial Intelligence: ECAI 2000*, IOS Press, Berlin, pp. 219-223.
- Guarino, N.** (1997a). *Some organizing principles for a unified top-level ontology*. AAAI Spring Symposium on Ontological Engineering, 57-63.
- Guarino, N.** (1997b). *Understanding, building and using ontologies*. International J. Human-Computer Studies, 46, 293-310.
- Guéuard, H.** (2002). *Anatomie et physiologie humaines*. Livre : 3^o édition.
- Haut, H.** (2003). *Le e-learning Un nouvel espace d'apprentissage*, TECHN 24 Publication technique de la SmalS-MvM, 04.
- Hocine, N., Gouaïch, A., Di Loreto, I., & Abrouk, L.** (2011). *Etat de l'art des techniques d'adaptation dans les jeux ludiques et sérieux*. Manuscrit auteur, publié dans 'numéro spécial RIA jeux sérieux 1 (2011) 30'. DOI : 10.3166/ria.25.253-280.
- Hong, H., & Kinshuk.** (2004). *Adaptation to student Learning styles in Web Based Educational Systems*. In L. Cantoni & C.McLoughlin (Eds.) *Proceeding of ED-Media 2004- World conference on Educational Multimedia, Hypermedia & telecommunications*, June 21- 26 2004, Lugano, Sxitzerland, USA : AACE,491-496 (ISBN 1-880094-53-3).
- HP** : 'Hewlett-Packard', *Labs Semantic Web Research*. En ligne : <http://www.hpl.hp.com/>
- Hyperwave.** (2013). Le site officiel. *L'architecture de la plate-forme Hyperwave*. En ligne : <http://www.hyperwave.com/e/>

- IAAT** : Institut Atlantique d'Aménagement des Territoires, Par le centre de documentation de l'IAAT. (2005). *Guide méthodologique du travail en commun*. 1-128.
- Inaba, A., Supnithi, T., Ikeda, M., Mizoguchi, R., & Toyoda, J.** (2000). *An Overview of Learning Goal Ontology*. Proc. of ECAI2000 Workshop on Analysis and Modelling of Collaborative Learning Interactions, 23-30.
- IRDES** : Institut de Recherche et Documentation en Economie de la Santé (2008). *Sources d'information et méthodologie de recherche documentaire*. Service documentation Irdes – Marie-Odile SAFON, 1-23.
- IRDES** : Institut de Recherche et Documentation en Economie de la Santé (2012). *Sources d'information et méthodologie de recherche documentaire*. Pôle documentation de l'Irdes – Marie-Odile Safon, 1-27.
- IWS** : Internet World Stats. (2012). *INTERNET USAGE STATISTICS*, The Internet Big Picture, World Internet Users and Population Stats. En ligne : <http://www.internetworldstats.com/stats.htm>
- JEA** : Java Excel API. (2011a). *A Java library for reading/writing Excel*. En ligne : 'http://sourceforge.net/projects/jexcelapi/files/'
- JEA** : Java Excel API. (2011b). *A Java API to read, write, and modify Excel spreadsheets*. En ligne : <http://jexcelapi.sourceforge.net/>
- JENA**. (2011a). *A Semantic Web Framework for Java*. En ligne : <http://jena.apache.org/>
- JENA**. (2011b). *Downloads*. En ligne : <http://jena.apache.org/download/index.html>
- Jin, L., Chen, W., Hayashi, Y., Ikeda, M., Mizoguchi, R., Takaoka, Y., & Ohta, M.** (1999). *An Ontology Aware Authoring Tool – Functional structure and guidance generation*. Paper presented at the AIED1999, Le Mans, France.
- Jovanovic, J., Gasevic, D., Knight, C., & Richards, G.** (2007). *Ontologies for effective use of context in e-learning settings*. Educational Technology & Society, 10 (3), 47–59.
- Karsenti, T., Savoie-zajc, L., & Larose, F.** (2001). *Les futurs enseignants confrontés aux TIC: changements dans l'attitude, la motivation et les pratiques pédagogiques, Education et Francophonie*.
- Khadir, M-T.** (2007). *Les ontologies concepts, méthodes et outils*, Support de cours LMD 3eme Année Gestion des connaissances, Annaba – Algérie.
- Khelil, M-A.** (1994). *Atlas d'anatomie*. Livre, Tlemcen – Algérie.
- Kolb.** (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Prentice Hall, Upper Saddle River.
- Kruse, K.** (2004). *The Benefits and Drawbacks of e-Learning*. Retrieved May 07, 2012, from http://www.corebiztechnology.com/software_article_elearning_d.htm

- Lee, Y.-H., Hsieh, Y.-C., & Hsu, C.-N.** (2011). *Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems*. *Educational Technology & Society*, 14(4), 124–137.
- Lee, M., Tsai, K., & Wang, T.** (2008). *A practical ontology query expansion algorithm for semantic-aware learning objects retrieval*. *Computers & Education*, 50(4), 1240–1257.
- Lérignier, R., & Marcheix, D.** (2013). *Domaine : Le traitement de l'information juridique, Module : Rechercher et utiliser des ressources d'information et de documentation juridique, Chapitre : Trois grandes catégories d'outils de recherche*.
- Liegle, J., & Janicki, T.** (2006). *The effect of learning styles on the navigation needs of Web-based learners*. *Computers in Human Behavior*, 22 (5), 885–898.
- Lindberg D. A. B., & Humphrey B. L., & McCray A. T.** (1993). *The Unified Medical Language System. Methods of Information in Medicine*. p. 81-91.
- Lu, E., & Hsieh, C.** (2008). *A relation metadata extension for SCORM content aggregation model*. *Computer Standards & Interfaces*, 31(5), 1028–1035.
- Maedche, A.** (2002). *Ontology Learning for the Semantic Web*. Boston: Kluwer Academic Publishers.
- Magoulas, G., Papanikolaou, K., & Grigoriadou, M.** (2003). *Adaptive Web-based learning: Accommodating individual differences through systems adaptation*. *British Journal of Educational Technology*, 34 (4), 511–527.
- Meniet, H.** (2007). *NetBeans*, En ligne : <http://www.trustonme.net/didactels/82.html>
- MIC : Bibliothèque Michel Serres.** (2007). *Pour une meilleure utilisation des moteurs de recherche*, Ecole Centrale de Lyon – France.
- Mizoguchi, R., & Bourdeau, J.** (2000). *Using Ontological Engineering to Overcome Common AI-ED Problems*. *International Journal of Artificial Intelligence and Education*, vol.11(Special Issue on AIED 2000), 107-121.
- Mizoguchi, R., & Ikeda, M.** (1996). *Towards Ontological Engineering (AI-TR-96-1)*. Osaka: ISIR, Osaka University.
- Mizoguchi, R.** (1998, June, 1998). *A Step Towards Ontological Engineering*. Paper presented at the 12th National Conference on AI of JSAI.
- Moisuc, B.** (2006). *Adaptation dans les systems d'information spatio-temporelle interactifs*, INFORSID 2006, 1er-3 juin, Hammamet, Tunisie.
- Monova-Zheleva, M.** (2005). *Adaptive learning in Web-based educational environments*. *Cybernetics and Information Technologies*, 5 (1), 44–55.
- Mumford, A., & Honey, P.** (1992). *Questions and answers on Learning Styles Questionnaire*. In *Industrial and Commercial Training*, Volume 24 (7), 10-13.

- Nesbit, J.C., Belfer, K., & Leacock, T.L.** (2004). *LORI 1.5: Learning Object Review Instrument, User manual*. Retrieved March 27, 2012 from <http://www.transplantedgoose.net/gradstudies/educ892/LORI1.5.pdf>
- NetBeans.** (2013). *Documentation Training & Support*. En ligne : <http://www.netbeans.org/kb/index.html>
- Nguyen, L., & Do, P.** (2008). *Learner Model in Adaptive Learning*. International Journal of World Academy of Science, Engineering and Technology WASET, 45 (70), 395–400.
- Olfman, L., & Mandviwalla, M.** (1994). *Conceptual versus procedural software training for graphical user interfaces: A longitudinal field experiment*. MIS Quarterly, 18(4), 405–426.
- Papanikolaou, K., Grigoriadou, M., Magoulas, G., & Kornilakis, H.** (2002). *Towards new forms of knowledge communication: The adaptive dimension of a Webbased learning environment*. Computers & Education, 39(4), 333–360.
- Peraya, D.** (2005). *La formation à distance : un dispositif de formation et de communication médiatisées. Une approche des processus de médiatisation et de médiation*, Genève – Suisse.
- Peterson, E. R., Rayner, S. G. & Armstrong, S. J.** (2009). *Researching the psychology of cognitive style and learning style: Is there really a future?*. Learning and Individual Differences, 19(4), 518–523.
- Pouilly, P.** (2007). *Recherche documentaire sur Internet*. Académie Dijon, 1-27.
- Psyché, V., Mendes, O., & Bourdeau, J.** (2003). *Apport de l'ingénierie ontologique aux environnements de formation à distance Contribution of Ontological Engineering to Distance Learning Environments*, publication francophone – STICEF, 1-32.
- PUS** : Protégé de l'Université Stanford. (2013a). *download protégé*. En ligne : <http://protege.stanford.edu/download/registered.html>, Etats-Unis.
- PUS** : Protégé de l'Université Stanford. (2013b). *user documentation*. En ligne : <http://protege.stanford.edu/doc/users.html#presentations>, Etats-Unis.
- PUS** : Protégé de l'Université Stanford. (2013c). *What is Protégé-2000?*. En ligne : http://protege.stanford.edu/doc/users_guide/, Etats-Unis.
- Rey-Lopez, M., Diaz-Redondo, R., Fernandez-Vilas, A., Pazos-Arias, J., Garcia-Duque, J., Gil-Solla, A., et al.** (2009). *An extension to the ADL SCORM standard to support adaptivity: The T-learning case study*. Computer Standards and Interfaces, 31(2), 309–318.
- Rézeau, J.** (2006). *Les outils de recherche sur Internet et sur le Web*. <http://joseph.rezeau.pagesperso-orange.fr/tutoriels/recherchesurinternet/outilsRecherche.htm>
- Riffaut, C., & et Rostaing, H.** (2002). *Ingénierie d'une formation à distance en enseignement supérieur*, CRRM Marseille - France.

- Rodica, M., & Anca, A.** (2008). *Knowledge Management in E-Learning Systems*. Revista Informatica Economica, 2 (46), 60–65.
- Rosmalen, P., Vogten, H., Van Es, R., Passier, H., Poelmans, P., & Koper, K.** (2006). *Authoring a full life cycle model in standards-based adaptive e-learning*. Educational Technology & Society, 9(1), 72-83.
- Roussel, G.** (2009) *API - Interface de programmation d'applications - Applications Programming Interface*, 2009.
- Ruiz, M., Diaz, M., Soler, F., & Perez, J.** (2008). *Adaptation in current e-learning systems*. Computer Standards & Interfaces, 30(1), 62–70.
- Sampson, D., Karagiannidis, C., & Cardinali, F.** (2002). *An architecture for Web-based e-learning promoting re-usable adaptive educational e-content*. Educational Technology & Society, 5(4), 27–37.
- Santally, M. I.** (2005). *A learning object Approach to Personalized Web-Based Instruction*. European Journal of Open, Distance and E-learning (EURODL).
- Scapin, D.L. & Bastien, J.M.C.** (1997). *Ergonomic criteria for evaluating the ergonomic quality of interactive systems*. Behaviour & Information Technology, 6(4-5), 220-231.
- Senan, M.** (2008). *Etude des méthodes de la recherche d'information et de l'indexation sur les documents électronique : Cas de la langue Arabe*, Doctorate thesis, (pp.21–22), Paris, France.
- Serres, A.** (2011). *SÉLECTION DE RESSOURCES SUR LES OUTILS DE RECHERCHE*. URFIST : Unité Régionale de Formation à l'Information Scientifique et Technique de Bretagne et des Pays de la Loire.
- Shih, W., Yang, C., & Tseng, S.** (2009). *Ontology-based content organization and retrieval for SCORM-compliant teaching materials in data grids*. Future Generation Computer Systems, 25(6), 687–694.
- Sowa, J.** (1995). *Distinction, combination, and constraints*. Proc. IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing.
- Stach, N., Cristea, A., & De Bra, P.** (2004). *Authoring of learning styles in adaptive hypermedia*. In Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters(pp. 114–123), New York, NY, USA.
- Stephanidis, C., Paramythis, A., Sfyarakis, M., Stergiou, A., Maou, N., Leventis, A., Paparoulis, G., & Karagiannidis, C.** (1998). *Adaptable and Adaptive User Interfaces for Disabled Users in AVANTI Project*. 5th International Conference on Intelligence in Services and Networks (IS&N '98), "Technology for Ubiquitous Telecommunications Services", Antwerp, Belgium.
- Surjono, H.** (2009). *The Development of an Adaptive E-Learning System Based on The E Learning Style Diversity of Visual-Auditory-Kinesthetic*. In International Seminar on ICT for Education, Yogyakarta State University, Indonesia, 13–14 February

- Ubuntu** : Communauté francophone des utilisateurs d'Ubuntu. (2012). *Netbeans*.
- UNJF/C2i** : Université Numérique Juridique Francophone / Certificat Informatique et Internet. (2010). *Le traitement de l'information juridique, Rechercher et utiliser des ressources d'information et de documentation juridique*.
- Uschold, M., & Grüninger, M.** (1996). *Ontologies: Principles, Methods and Applications*. Knowledge Engineering Review, 11(2).
- Uschold, M., & King, M.** (1995). *Towards a Methodology for Building Ontologies*. Paper presented at the Workshop on Basic Ontological Issues in Knowledge Sharing.
- UVMaF** : Université Virtuelle de Maïeutique Francophone (2011). *La recherche documentaire*. Comité éditorial de l'UVMaF, Université Médicale Virtuelle Francophone, 1-22.
- Van-Heijst, G., Schreiber, A., & Wielinga, B. J.** (1997). *Using Explicit Ontologies in KBS Development*. International Journal of Human and Computer Studies /Knowledge Acquisition, 46(2/3), 183-292.
- Vanwelkenhuysen, J., & Mizoguchi, R.** (1994). *Maintaining the workplace context in a knowledge level analysis*. Paper presented at the Proc. of JKAW'94, Hatoyama, Japan.
- Vanwelkenhuysen et Mizoguchi,** (1995) Vanwelkenhuysen, J., & Mizoguchi, R. (1995). *Workplace-Adapted Behaviors: Lessons Learned for Knowledge Reuse*. Paper presented at the KB&KS '95.
- Verbert, K., Gasevic, D., Jovanovic, J., & Duval, E.** (2005). *Ontology-based learning content repurposing*. In Special interest tracks and posters of the 14th international conference on World Wide Web,(pp. 1140–1141), ACM Press, Chiba, Japan.
- Villanova-Oliver, M.** (2002). *Adaptabilité dans les systèmes d'information sur le web : modélisation et mise en oeuvre de l'accès progressif*. Thèse de Doctorat à l'institut national polytechnique de Grenoble, France.
- Wang, M., Ran, W., Liao, J., & Yang, S.J.H.** (2010). *A Performance-Oriented Approach to E-Learning in the Workplace*. Educational Technology & Society, 13(4), 167-179.
- Wang, M., Peng, J., Cheng, B., Zhou, H., & Liu, J.** (2011). *Knowledge Visualization for Self-Regulated Learning*. Educational Technology & Society, 14(3), 28–42.
- Wang, H.-C., & Hsu, C.-W.** (2006). *Teaching-material design center: An ontology based system for customizing reusable e-materials*. Computers & Education, 46(4), 458–470.
- Wang, T., Wang, K., & Huang, Y.** (2008). *Using a style-based ant colony system for adaptive learning*. Expert Systems with Applications, 34(4), 2449–2464.
- Wielinga, B., & Schreiber, A.** (1993). *Reusable and sharable knowledge bases: A European perspective*. Paper presented at the KB & KS'93, Tokyo.

- WIKIPEDIA.** (2013). *Filtrage d'Internet.* En ligne : http://fr.wikipedia.org/wiki/Filtrage_d%27Internet
- WIKIVERSITY.** (2013). *Filtrage des informations : Définition.* En ligne : http://fr.wikiversity.org/wiki/Filtrage_des_informations/D%C3%A9finition
- Yang, Y., & Wu, C.** (2009). *An attribute-based ant colony system for adaptive learning object recommendation.* *Expert Systems with Applications*, 36(2), 3034–3047.
- Yang, S.** (2006). *Context aware ubiquitous learning environments for peer-to-peer collaborative learning.* *Educational Technology & Society*, 9(1), 188–201.
- Yessad, A. Faron-Zucker, C. Dien-Kuntz, R. Laskri, M. T.** (2008). *Adaptive Learning Organizer for Web-based Education.* *International Journal of Web-Based Learning and Teaching Technologies*, 3(4), 57–73.
- Zaidi, S.** (2008). *Les ontologies, Cours de Master 2 IC : ingénierie des connaissances, Annaba – Algérie.*
- Zitko, B., Stankov, S., Rosic, M., & Grubisic, A.** (2009). *Dynamic test generation over ontology-based knowledge representation in authoring shell.* *Expert Systems with Application*, 36 (4), 8185–8196.

Références Personnelles

▪ Publications Internationales

Mohammed, C.; Tayeb, L.M. (2010). *Automatic construction of an on-line learning domain*. IEEE CONFERENCES, (ICMWI'2010 : International Conference on Machine and Web Intelligence, October 3-5, 2010, Algiers, Algeria), ISBN: 978-1-4244-8608-3, DOI: [10.1109/ICMWI.2010.5648199](https://doi.org/10.1109/ICMWI.2010.5648199), pp. 439-443, November 2010.

Chaoui, M., & Laskri, M.T. (2011a). *New method of finding information on the Web in unstructured information resources for educational use by learners*. International Journal of Research and Reviews in Computer Science (IJRRCS), 2(1), 33–39, Science Academy Publisher, United Kingdom.

Chaoui, M., & Laskri, M-T. (2011b). *Towards the Creation of Adaptive Content from Web Resources in an E-Learning Platform to Learners Profiles*. International Journal of World Academy of Science, Engineering and Technology WASET, 77 (27), 157–162.

Chaoui, M., & Laskri, M-T. (2012). *New Web Tool to Create Educational and Adaptive Courses in an E-Learning Platform Based Fusion of Web Resources*. CEUR-WS Workshop Proceedings, (ICWIT'12 : The 4th International Conference on Web and Information Technologies, Sidi Bel Abbes, Algeria, April 29-30, 2012), published on CEUR-WS.org, Vol-867, urn:nbn:de:0074-867-1, ISSN 1613-0073, pp. 306-311, July 05, 2012.

Chaoui, M., & Laskri, M. T. (2013). *Proposition and Organization of an Adaptive Learning Domain based on Fusion from the Web*. Educational Technology & Society, 16 (1), 118–132.

• Publications Nationales

Mohammed Chaoui. *LA PLATE-FORME MOODLE*. JCEL2011, Workshop on: COURS EN LIGNE “Course online”. University center of El-Oued, November 20, 2011.

Mohammed Chaoui, Mohamed Tayeb Laskri. *The use of Web resources for creating educational and adaptive content to learners in an e-Learning platform*. Events with acts: COSI2011, 8th Symposium on Optimization and Information Systems, Guelma, Algeria, April 24-28, 2011.

Mohammed Chaoui, Mohamed Tayeb Laskri. *Vers une nouvelle extension pour les éditeurs LCMS dans un système e-Learning*. Events with acts: STIC'11, 2nd Doctorial in Sciences & Technologies of Information and Communication, Tébessa, Algeria, April 20-21, 2011.

• International Project

Industry Gateway (international database) with a certificate number : 003352 to acknowledge the publication of our work and our contribution to Industry Gateway's Database “*Towards the creation of adaptive content from Web resources in an e-Learning platform to learners profiles*” (003352), 01/07/2011.

Liste des Abréviations

Abréviation	Description Textuelle
AICC	Aviation Industry CBT Committee
ALD	Adaptive Learning Domain
API	Application Programming Interface
API SAX	Simple API for XML
ARCHONTE	ARCHitecture for ONTological Elaborating
ARIADNE	Alliance of Remote Instructional Authoring and Distribution Networks for Europe
BDD	Base De Données
BNF	Base de Nouveaux Fragments
CDDL	Common Development and Distribution License
CGI	Common Gateway Interface
DA	Domaine d'Apprentissage
DAA	Domaine d'Apprentissage Adaptatif
DAML	DARPA Agent Markup Language
DBRS	Distance Basée sur des Règles Sémantiques
DBRS_D	Formule de la Distance dans la méthode DBRS
DBRS_M	Formule de la Moyenne dans la méthode DBRS
DBRS_{PA}	Formule de la Précision d'une Association dans la méthode DBRS
DBSR	Distance Based on Semantic Rules
DCMI	Dublin Core Metadata Initiative
DOM	Document Object Model
dotLRN	dot LRN Learn, Research, Network
DP	Degré de Pertinence
DR	Degree of Relevance

EDI	Electronic Data Interchange
FLOGIC	Frame LOGIC
FSF	Fusion des Sous Fragments
FTP	File Transfer Protocol
GPL v2	General Public Licence Version 2
HP	Hewlett-Packard
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protocol
IA	Intelligence Artificielle
ICT	Information and Communication Technologies
IDE	Integrated Development Environment
IETF	Internet Engineering Task Force
IMAT	Integrating Manuals And Training
Java IDL	Java Interface Definition Language
Java RMI	Java Remote Method Invocation
KB	Knowledge Base
KIF	Knowledge Interchange Format
LCMS	Learning Content Management System
LD	Learning Domain
LMS	Learning Management System
LOM	Learning Object Metadata
LORI	Learning Object Review Instrument
N3	Notation3 pour RDF
NFDB	New Fragments DataBase
NTIC	Nouvelles Technologies de l'Information et de la Communication
OCML	Operational Conceptual Modelling Language

ODE	Ontology Design Environment
OIL	Ontology Inference Layer or Ontology Interchange Language
OilEd	Oil Editor
OKBC	Open Knowledge Base Connectivity
OLAT	Online Learning And Training
OLMS	Open Learning Management System
OntoEdit	Ontology Editor
OWL	Ontology Web Language
OWL DL	Description Logic of OWL file
PA	Précision d'une Association
PC	Personal Computer
PDF	Portable Document Format
PHP	Hypertext Preprocessor
PrOALDoF-Web	Proposition and Organization of an Adaptive Learning Domain based Fusion from the Web
QBLS	Question Based Learning System
RDF	Ressource Description Framework
RDFS	RDF Shema
RS	Règles Sémantiques
SCD	Service commun de la documentation
SCORM	Sharable Content Object Reference Model
SF	Sous Fragment
SOAP	Simple Object Access Protocol
SPARQL	SPARQL Protocol and RDF Query Language
TCP/IP	Transmission Control Protocol/Internet Protocol
TIC	Technologies de l'Information et de la Communication

UMLS™	Unified Medical Language System
UNESCO	United Nations Educational, Scientific and Cultural Organization
URI	Universal Resource Identifiers
URL	Uniform Resource Locator
USC	University of Southern California
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	eXtensible Markup Language