Table des matières

Introduction Générale **10**

Chapitre I : Les ontologies **13**

I .1. Introduction 14

I .2. Historique 14

I .3. Définition 15

I .4. Rôle des ontologies 15

I .5. Les constituants d'une ontologie 16

I .5.1. Les classes/ les concepts 17

I .5.2. Les relations et les fonctions 18

I .5.3. Les axiomes 18

I .5.4. Les instances 18

I .6. Le cycle de vie des ontologies 18

I .6.1. Besoins et évaluation 19

I .6.2. Conception etévolution 19

I .6.3. Diffusion 20

I .6.4. Utilisation 20

I .6.5. Gestion 20

I .7. Classification des ontologies 20

I .7.1. Typologie selon la richesse de la structure interne des ontologies 21

a) Le vocabulaire contrôlé 21

b) Le glossaire 21

c) Le thésaurus 21

d) La hiérarchie informelle (is-a) 22

e) La hiérarchie formelle (is-a) 22

f) La hiérarchie formelle (is-a) avec instances du domaine 22

g) Le frame 23

h) Les ontologies avec restrictions de valeur 23

i) Les ontologies avec contraintes logiques 23

I .7.2. Typologie selon l'objet de conceptualisation 23

a) Les ontologiesde représentation de connaissances 23

b) Les ontologies générales ou communes ou génériques 23

c) Les ontologies de niveau supérieur ou de hautniveau 23

d) Les ontologies du domaine 24

e) Les ontologies de tâches 24
f) Les ontologies d'application 24
I .7.3. Typologie selon le niveau de granularité 24
a) La granularité fine 24
b) La granularité large 25
I .7.4. Typologie selon le niveau de formalisation utilisé 25
I .8. Méthodes et Méthodologies de Construction d'ontologie 26
I .8.1. Le processus de développement d'ontologie 26
I.8.2. Les méthodes et méthodologies de développement des ontologies 28
a) La méthodologie METHONTOLOGY 28
b) La méthodologie On-To-Knowledge 29
c) La méthode Ontology Development 101 29
d) Les autres méthodes de développement d'ontologies 29
I .9. Domaine d'application des ontologies 31
I .9.1. Système d'information 31
I .9.2. Web sémantique 32
I .10. Conclusion 32

Chapitre II : Les langages et les outils d'ontologie 33
II .1. Introduction 34
II .2. Les langages de description des ontologies 34
II .2.1. Les langages d'ontologies traditionnels 34
II .2.2. Les langages d'ontologie web standard ou baséssur XML 35
II .3. Les outils de construction d'ontologie 37
II .3.1. Les outils dépendants du formalisme de représentation 37
a. Ontolingua 37
b. OntoSaurus 37
c. WebOnto 38
d. OilEd 38
II .3.2. Les outils indépendants de formalisme de représentation 38
a. Protégé2000 38
b. ODE et WebOde 38
c. OntoEdit 38
II .4. Conclusion 38

Chapitre III : Etat De l'art

111.1. Introduction..... 41
111.2. Terminologies..... 41
111.3. Ontologies Vs Terminologies..... 41
111.4. Représentation des ontologies..... 42
III .4.1. RDF et RDF(S)..... 43
III .4.2. OWL..... 44
a. OWL Full..... 44
b. OWL DL..... 44
c. OWL Lite..... 44
111.5. Construction et édition d'ontologies..... 45
III.5.1. Les méthodes et méthodologies pour la construction d'ontologies en partant de zéro.... 46
III .5.2. Les méthodes pour la réingénierie d'ontologies..... 46
III .5.3. Les méthodes pour la construction coopérative d'ontologies..... 46
111.6. Enrichissement, maintenance et évolution d'ontologies..... 47
111.7. Exemples d'ontologies et de classifications dans le domaine médical..... 48
III .7.1. UMLS..... 49
III .7.2. SNOMED CT..... 49
III .7.3. MeSH..... 49
III .7.4. Galen..... 49
III .7.5. Digital Anatomist : Foundational Model of Anatomy (FMA)..... 50
111.8. Conclusion..... 50

Chapitre IV : Application..... 50
IV.1. Introduction..... 51
IV.2. Outils et langages utilisés..... 51
IV.2.1. NetBeans..... 51
IV.2.2. Jena..... 51
IV.2.3. JAVA..... 52
IV.3. Interfaces..... 52
IV.4. Conclusion..... 59

Conclusion Générale..... 60

Référence Bibliographique..... 62

Liste de figure

Figure I. 1: Le triangle sémantique (Ogden et Richards, 1923). [7]
Figure I. 2: cycle de vie d'une ontologie. [7]
Figure I. 3: Classification de Lassila et Mc Guinness. [12]
Figure I. 4: Processus de développement d'ontologie de METHONTOLOGY. [25]

Figure I. 5: Pyramide du web sémantique. [41]

Figure II. 1: Les langages d'ontologies traditionnels [28]
Figure II. 2: Les langages d'ontologie basés sur XML [28]

Figure IV. 1 : Page d'accueil jena
Figure IV.2 : Copie d'écran de l'interface principale
Figure IV. 3 : Copie d'écran de l'ouverture d'ontologie
Figure IV.4 : Copie d'écran de l'ontologie enregistrée
Figure IV. 5 : Copie d'écran d'ajout d'un patient
Figure IV. 6 : Copie d'écran d'ajout d'une maladie
Figure IV. 7 : Copie d'écran d'ajout d'un médicament
Figure IV. 8 : Copie d'écran des individus ajoutés
Figure IV. 9 : Copie d'écran des relations
Figure IV. 10 : Copie d'écran de la relation ajoutée
Figure IV. 11 : Copie d'écran d'ajout d'un Nom
Figure IV. 12 : Copie d'écran d'ajout d'un Prénom
Figure IV. 13 : Copie d'écran d'ajout d'une date de naissance
Figure IV. 14 : Copie d'écran d'ajout d'un code
Figure IV. 15 : Copie d'écran des types de propriété de donnée
Figure IV. 16 : Copie d'écran d'ajout d'une propriété de donnée DateN
Figure IV. 17 : Copie d'écran du nouveau graphe
Figure IV. 18 : Copie d'écran de Relation_Individu
Figure IV. 19 : Copie d'écran de nouveau graphe « Omar A pour maladie Migraine »
Figure IV. 20 : Copie d'écran d'ajout de concept « B-sanguin »
Figure IV. 21 : Copie d'écran d'ajout d'une nouvelle classe
Figure IV. 22 : Copie d'écran d'ajout de concept « Diabète »
Figure IV. 23 : Copie d'écran d'un ajout d'un nouveau concept
Figure IV. 24 : Copie d'écran de suppression de concept « B-Sanguin »
Figure IV. 25 : Copie d'écran du nouveau graphe

Liste des abréviations

G.O.M	Gestion des ontologies Médicales.
OWL	Web OntologyLanguage.
IEEE	Institute of Electrical and Electronics Engineers.
ODE	Ontology Design Environment.
FIPA	Foundation for Intelligent Physical Agents.
URI	Universal Resource Identifier. extensible Markup Language.
XML	World Wide Web Consortium.
W3C	Resource Description Framework.
RDF	Resource Description Framework Schema.
RDFS	SPARQL Protocol and RDF Query Language.
SPARQL	Rule Interchange Format.
RIF	Hypertext Markup Language.
HTML	Knowledge Interchange Format.
KIF	Operational Conceptual Modeling Language.
OCML	Frame Logic.
Flogic	Open Knowledge Base Connectivity.
OKBC	Simple HTML Ontology Extension.
SHOE	XML-based Ontology exchange Language.
XOL	Ontology Interchange Language and Ontology Inference Layer.
OIL	DARPA Agent Markup Language.
DAML+OIL	Ontology Web Language.
OWL	Oil Editor.
OilEd	Ontology Editor.
OntoEdit	les Logiques de Description.
DLs	Uniform Resource Identifier.
URI	OWL Description Logic.
OWL DL	National Library of Medicine.
NLM	Unified medical Language System.
UMLS	Generalised Architecture for Languages, Encyclopaedias and Nomenclatures in Medicine
GALEN	

CUI	Concept Unique Identifier.
SUI	String Unique Identifier.
LUI	Lexique Unique Identifier.
SNOMED-CT	Systematized Nomenclature of Medicine, Clinical Terms SNOMED RT Reference Terminology.
MeSH	Medical Subject Headings.
NLM	National Library of Medicine.
CORE	Galen Coding Reference.
GRAIL	GALEN ReprésentationAnd Integration Language.
FMA	Foundational Model of Anatomy.
NBH	Neuronames Brain Hierarchy.
CDDL	Common Dévelopmentand Distribution License.
IDE	Integrated Development Environement.

Introduction Générale

La transmission des informations des activités médicales des patients par les établissements hospitaliers est devenue obligatoire. Ces informations sous forme de résumés standardisés concernent les diagnostics principaux et secondaires et sont effectuées à partir de la classification internationale des maladies. En utilisant un thésaurus de spécialité, proposés par des sociétés prestataires de services, les médecins réalisent leur codage le plus souvent manuellement. Ces codages se révèlent dans la majorité des cas ambigus et non exhaustifs. Les systèmes informatiques de codage disponibles sur le marché utilisent généralement des thésaurus et de portée limitée. Pour pallier à cette insuffisance il est nécessaire de décrire la sémantique des objets et de l'organisation du domaine médical, et ceci dans le but de développer des modèles conceptuels qui ne sont ni contextuel ni ambiguë, et dont le sens est inscrit dans la structure même du modèle. Cette modélisation est appelé «ontologie». L'un des principaux objectifs de la recherche est que le développement de ces ressources ontologiques garantisse des outils puissants, fiables, et maintenables pour le codage. La problématique à laquelle nous nous attaquons est donc la gestion de ces ontologies.

Les ontologies médicales ou autres, sont construites en étroite collaboration avec les utilisateurs, avec lesquels sont mis en place des procédures d'évaluation des ressources dans leur contexte d'usage.

La recherche sur les ontologies qui a débuté vers les années 90 du siècle dernier dans les domaines de l'intelligence artificielle, de l'ingénierie des connaissances, et de la représentation des connaissances, a vu son champ d'application s'élargir de façon considérable, et fait partie des axes de la recherche actuelle, en particulier dans le domaine de la modélisation des systèmes d'information où les recherches utilisant les ontologies sont de plus en plus nombreuses. Les ontologies sont des systèmes formels dont l'objectif est de représenter les connaissances d'un domaine au moyen de concepts, définis et structurés les uns par rapport aux autres. La représentation ontologique des connaissances assure le maintien de la cohérence des axiomes et de l'intégrité du système, ainsi que l'évolution de la représentation sans changer la structure.

Pour l'implémentation de notre application nous avons utilisé l'environnement de développement intégré NetBeans IDE avec les packages Jena et le langage Java.

Notre projet de fin d'étude et structurer en quatre chapitres :

- Après cette introduction le premier chapitre est consacré aux ontologies, leurs définitions, historiques cycles de vie etc...
- Le deuxième chapitre s'intéresse aux différents outils et langages d'ontologies.
- Dans le troisième chapitre nous présentons un état de l'art sur les ontologies.

Rapport-Gratuit.com

- Le quatrième chapitre détaille à l’implémentation de l’ontologie médicale et l’application.
- Et nous terminons par une conclusion où nous rappellerons les défèrent parties clés de notre travail.

Chapitre I

Les ontologies

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

I .1. Introduction

Le développement et l'exploitation des connaissances en informatique a tellement évolué, qu'un des principaux objectifs de la recherche actuelle est de ne plus considérer l'ordinateur comme une boite noire, dans laquelle, sont stockées des informations, mais plutôt comme une machine intelligente avec laquelle on peut dialoguer et créer une coopération. Le système doit alors, avoir accès non seulement aux termes utilisés par l'être humain mais aussi à la sémantique qui leurs est associée, afin qu'une communication efficace soit possible.

La connaissance visée par ces ontologies est un thème de recherche dans divers axes tels que l'ingénierie des connaissances, la recherche d'information, les systèmes d'information coopératifs, l'intégration intelligente d'information, la gestion des connaissances ... Elles offrent une connaissance partagée sur un domaine qui peut être échangée entre des personnes et des systèmes hétérogènes. Elles ont été définies en intelligence artificielle afin de faciliter le partage des connaissances et leurs réutilisations. La définition explicite du concept ontologie soulève un questionnement qui tout à la fois d'ordre philosophique, cognitif et technique.

I .2. Historique

Une ontologie en informatique est un ensemble structuré de concepts permettant de donner un sens aux informations. Les concepts sont organisés dans un graphe dont les relations peuvent être:

- des relations sémantiques;
- des relations de composition et d'héritage (au sens objet)

L'objectif premier d'une ontologie est de modéliser un ensemble de connaissances dans un domaine donné.

À l'origine, l'ontologie est une notion philosophique, et on considérait que l'étude de l'ontologie était une partie de la métaphysique, qui s'intéresse à l'étude des propriétés de l'être, et par extension de l'existence. [70]

Le mot ontologie provient du grec « onto » qui signifie ce qui existe, comme l'être, et l'existant, et « logos » qui veut dire étude.

[71] John McCarthy a introduit l'ontologie en intelligence artificielle en 1980, par le principe que les concepteurs de systèmes intelligents fondés sur la logique devraient d'abord établir une base de données complète de ce qui existe, et ensuite utiliser ces données.

Les ontologies sont apparues en intelligence artificielle, comme réponses aux problématiques de représentation et de manipulation des connaissances au sein des systèmes informatiques.

I .3. Définition

Dans le cadre de l'intelligence artificielle, Neeches et ses collègues [1] furent les premiers à proposer une définition à savoir « une ontologie définit les termes et les relations de base de vocabulaire d'un domaine, ainsi que les règles qui indiquent comment combiner les termes et les relations de façon à pouvoir étendre le vocabulaire ».

En 1993, GRUBER [2] propose la définition suivante : « une ontologie est une spécification explicite d'une conceptualisation ».

En 1995, GUARINO [3] a modifié légèrement la définition de GRUBER, et la définit par: «une ontologie est une spécification partielle et formelle d'une conceptualisation».

En 1997, ces deux dernières définitions sont regroupées dans celle de BORST [4]

« Une ontologie est définit comme étant une spécification formelle d'une conceptualisation partagée».

- Conceptualisation : le modèle abstrait d'un phénomène du monde réel par identification des concepts clefs de ce phénomène.
- Explicite : le type de concepts utilisés et les contraintes liés à leur usage sont définis explicitement.
- Formelle : une ontologie doit être traduite en langage interprétable par une machine.
- Partagée : capture la connaissance consensuelle, c'est-à-dire non réservée à quelques individus, mais partagée par un groupe ou une communauté.

I .4. Rôle des ontologies

Les ontologies qui sont actuellement au centre de la recherche dans le domaine du Web sémantique ont été développées dans le but de représenter des connaissances, que les machines peuvent comprendre, peuvent manipuler la sémantique des informations. La construction d'ontologies qui nécessite une étude de la connaissance humaine, pour la mise en œuvre des systèmes de manipulation dépend des langages de représentation. L'étude des connaissances humaines et plus largement de l'esprit humain, qui rattache l'Intelligence Artificielle aux sciences humaines, et la création de programmes évolutifs reflétant le travail d'un groupe de chercheurs (artefacts) représentent les dimensions scientifiques et technique de la participation des ontologies dans le domaine de l'Intelligence Artificielle. Au fil des années, les méthodologies de construction d'ontologies et les outils de développement adéquats sont apparus.

Partant d'une pratique limitée voir artisanale, les méthodologies de construction d'ontologies et les outils de développement adéquats se sont développés, au fil des années mettant

à profit l'expérience et constitue actuellement une véritable ingénierie de construction et de gestion d'ontologie tout au long de leur cycle de vie.

Les ontologies représentent la composante logicielle qui s'intègre dans les systèmes d'information en leur fournissant un aspect sémantique, et de ce fait, leur champ d'application s'est considérablement élargie et couvre les systèmes conseillers comme les systèmes d'aide à la décision, les systèmes d'enseignement assisté par ordinateur " e-learning ", etc., les systèmes de résolution de problèmes et les systèmes de gestion de connaissances.

Parmi les projets d'actualité figure l'utilisation des ontologies qui peut procurer au Web une plate-forme de connaissances permettant, la recherche d'information des points de vue syntaxique et sémantique.

L'enjeu de l'effort engagé est de rendre les machines suffisamment sophistiquées pour qu'elles puissent intégrer le sens des informations, qu'elles ne font que manipuler formellement, à l'heure actuelle. Mais en attendant que des ordinateurs chargés d'ontologies et de connaissances nous soulagent en partie du travail de plus en plus lourd de gestion des informations dont le flot a tendance à nous submerger, où de nombreux problèmes théoriques et pratiques restent à résoudre.

I .5. Les constituants d'une ontologie

Les ontologies définissent, le sens des termes et les relations entre eux, elles leur fournissent un vocabulaire commun. Malgré que les ontologies peuvent être très différentes surtout au niveau du traitement de leurs composantes de base telles que les choses, les relations..., elles caractérisent un même univers.

I .5.1. Les classes/ les concepts

Les classes ou concepts, représentent des types d'objets. Elles, sont habituellement organisées en taxinomies soumis aux mécanismes d'héritage. Elles peuvent être structurées en une large taxinomie, avec un grand nombre de hiérarchisations ou sans taxinomie.

Les concepts sont utilisés dans leur sens large. Ils peuvent être de natures très différentes abstraits ou concrets, élémentaires (électrons) ou composés (atome), réels ou fictifs.

Il arrive que les définitions des ontologies aient été diluées, en ce sens que les taxinomies sont considérées comme des ontologies complètes. [5]

Un concept peut donc décrire une tâche, une fonction, une action, une stratégie, un processus de

raisonnement, etc.

Selon PINTO00 [6], Il est souvent fait référence aux concepts en tant qu'union de classes et d'instances, alors que chacun des constituants de l'ontologie est considéré comme un fragment de connaissance.

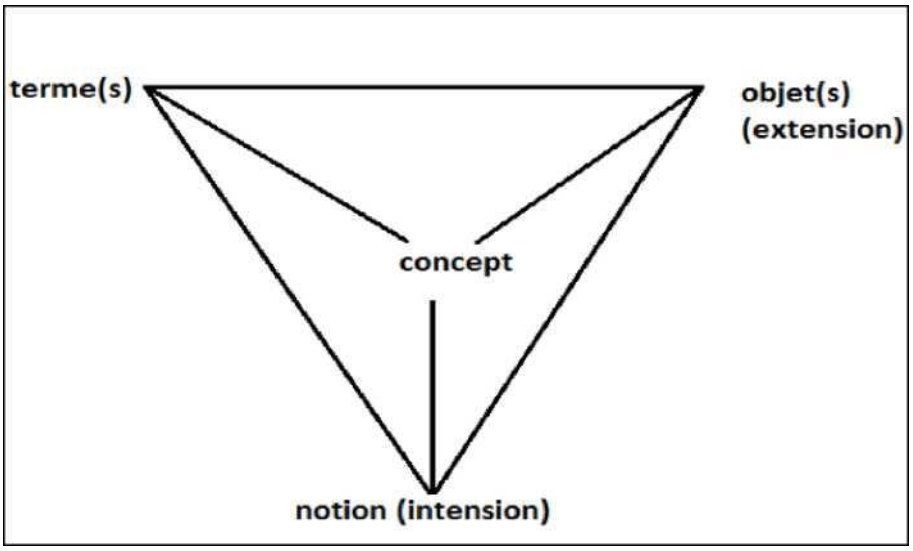

Figure I. 1 : Le triangle sémantique (Ogden et Richards, 1923). [7]

I .5.2. Les relations et les fonctions

L'interaction entre les notions d'un domaine définit les relations, On trouve des relations binaires, tertiaire.Certaines relations binaires entre des objets sont considérées comme des rôles comme le définit Borgida [8].

Les fonctions sont des cas particuliers de relations dans lesquelles un élément de la relation est défini à partir des autres éléments.

I .5.3. Les axiomes

Les axiomes désignent des vérités indémontrables qui doivent être admises. Ce sont des affirmations considérées comme évidentes sans preuve. Ils permettent de contraindre les valeurs de classes ou d'instances.

I .5.4. Les instances

Les instances représentent les éléments des concepts et des relations dans un domaine donné. [9]

I .6. Le cycle de vie des ontologies

Il existe plusieurs types d'activités liées aux ontologies ; on peut citer :

- Les activités de gestion ;
- Les activités de développement (spécification, conceptualisation, formalisation) ;
- Les activités transversales de support.

Un cycle de vie d'une ontologie débute du besoin qui se transforme en idée, la concrétisation de l'idée qui se traduit par la conception qui est diffusée pour son utilisation. Vient ensuite l'étape de l'évaluation qui donne naissance, le plus souvent à une étape d'évolution.et de maintenance du modèle. Une réévaluation de l'ontologie et des besoins devra se faire après chaque utilisation significative. [10]

L'ontologie peut être étendue et, si nécessaire, en partie reconstruite. La validation du modèle de connaissances est au centre du processus et se fait de manière itérative. [11]

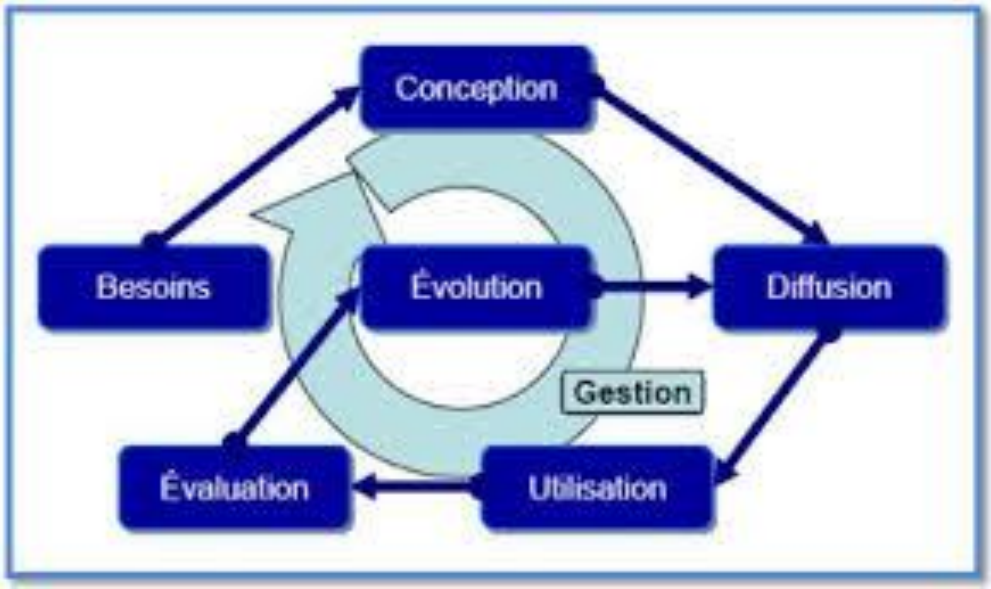

Figure I. 2: cycle de vie d'une ontologie. [7]

I .6.1. Besoins et évaluation

Dans l'utilisation d'une ontologie, des problèmes méthodologiques de recueil d'information, d'analyse et d'identification peuvent surgir. Afin d'éviter ces problèmes dans l'utilisation d'une ontologie, un état des lieux initial approfondi doit être élaboré dans l'étape de détection des besoins, car cette étape ne peut reposer sur des études précédentes ou des retours d'utilisation comme dans le cas de l'activité d'évaluation.

I .6.2. Conception et évolution

Les phases de conception initiale et d'évolution ont en commun un certain nombre de points :

- spécification des solutions ;
- conceptualisation et modélisation ;
- formalisation (logiques de description, graphes conceptuels, formalismes du web sémantique RDF, RDF(S) et OWL) ;
- intégration de ressources existantes ;
- implantation (graphes conceptuels, logiques de description).

Le choix de représentation et de conceptualisation faits dans l'ontologie représente un problème de conception et d'évolution. Notons aussi que l'évolution pose le problème de la maintenance de ce qui repose déjà sur l'ontologie. L'ontologie est à la fois un ensemble évolutif et un ensemble de primitives pour décrire des faits et des algorithmes sur ces faits. Ses changements donc, ont un impact direct sur tout ce qui a été construit sur la base de cette ontologie. Le maintien de la cohérence dans une ontologie est un des points clés dans son utilisation La maintenance de l'ontologie soulève donc des problèmes d'intégration technique et des problèmes d'intégration aux usages. Les domaines de la maintenance comme l'historique et la gestion des versions, la propagation des changements après modification, sont des questions importantes de la recherche actuelle.

I .6.3. Diffusion

Le déploiement et à la mise en place de l'ontologie interviennent dans la phase de diffusion, où les problèmes sont fortement contraints par l'architecture des solutions. Pour l'application Web on utilisera des technologies adéquates et pour le partage de fichiers, nous pouvons utiliser des architectures distribuées ou l'architectures Peer to Peer. Dans toutes les architectures dans les serveurs web, services web, Peer to Peer, etc. la distribution des ressources et leur hétérogénéité du point de vue syntaxiques, sémantiques, protocolaires, contextuelles, ou autres posent des problèmes de recherche sur l'interopérabilité et le passage à l'échelle (larges bases, optimisation d'inférences, propagation de requêtes, etc.).

I .6.4. Utilisation

Les activités reposant sur la disponibilité de l'ontologie, comme l'annotation des ressources, la résolution de requête, la déduction de connaissances, etc. constituent la phase d'utilisation.

Les problèmes de la conception des interactions avec l'utilisateur et de leur ergonomie concernant les interfaces dynamiques, les profils et contextes d'utilisation, sont posés dans toutes ces activités. Sur ce point, l'ontologie apporte à la fois de nouvelles solutions et de nouveaux problèmes.

I .6.5. Gestion

L'existence d'une activité permanente de gestion et planification est importante pour assurer une pérennité dans le travail de suivi et la politique globale pour la détection, la préparation l'évaluation des itérations d'un cycle et s'assurer que l'ensemble de ces causes améliorent les systèmes d'information.

I .7. Classification des ontologies

Les classifications des ontologies peuvent être faites suivant plusieurs paramètres. Parmi les plus répandues nous citons :

- La richesse de la structure interne des ontologies ;
- L'objet de conceptualisation ;
- Le niveau de granularité ;
- Le niveau de formalisation de la représentation des connaissances.

I .7.1. Typologie selon la richesse de la structure interne des ontologies

Lassila et Mc Guinness [12] proposent une classification des ontologies selon la richesse de leur structure interne. Cette classification consiste en un continuum allant du vocabulaire contrôlé à l'ontologie avec des contraintes logiques. Lorsque nous nous déplaçons tout le long du continuum, le sens des termes devient plus précis et le degré de formalisme augmente ce qui réduit l'ambiguïté. Ce critère est connu aussi dans d'autres travaux par le degré d'engagement sémantique qui correspond au niveau de spécification formelle permettant de restreindre l'interprétation de chaque concept et ainsi d'en donner la sémantique [13]. Autrement dit, le degré d'engagement sémantique fait en particulier référence au niveau sémantique des connaissances que l'ontologie représente.

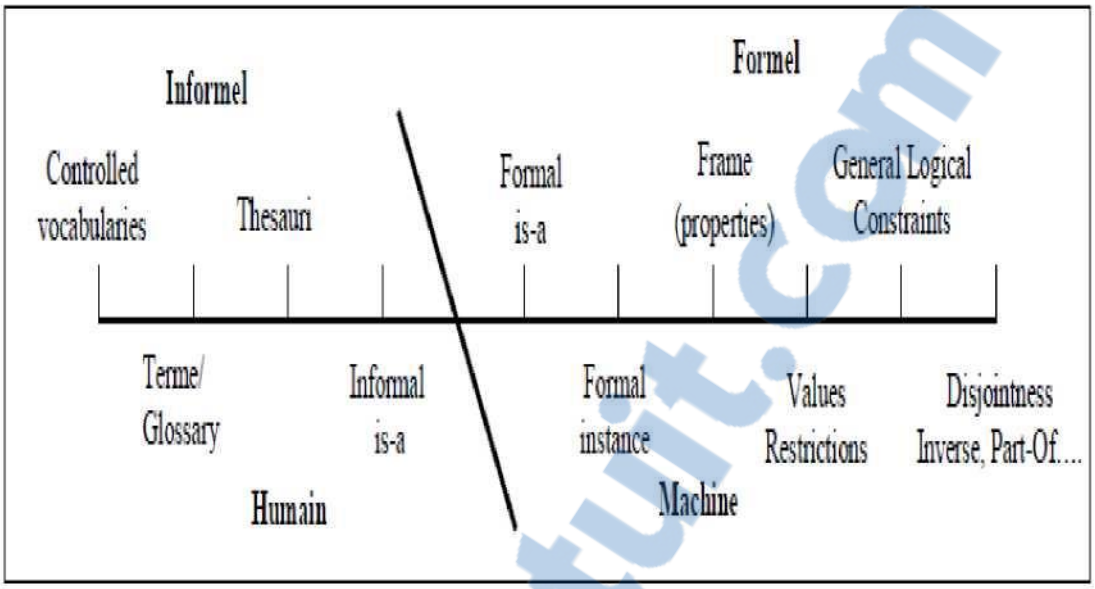

Figure I. 3: Classification de Lassila et Mc Guinness. [12]

Les principales catégories de cette classification sont:

a) Le vocabulaire contrôlé

C'est une liste finie de termes définis par un groupe de personnes ou une communauté. La signification des termes n'est pas forcément définie et il n'y a pas d'organisation logique entre les termes. Les catalogues sont des exemples de cette catégorie.

b) Le glossaire

C'est une liste de termes avec leurs significations spécifiées en langage naturel. Cette représentation apporte plus d'informations car une personne peut lire la définition, cependant elle n'est pas interprétable par l'ordinateur.

c) Le thésaurus

Il fournit une sémantique supplémentaire entre les termes tels que la relation de synonymie. Mais, il n'offre pas la structure hiérarchique explicite.

d) La hiérarchie informelle (is-a)

C'est une hiérarchie où la notion vague de généralisation et de spécialisation est fournie bien que ce ne soit pas une hiérarchie qui correspond à la stricte notion de subsomption. Ce qui revient à dire que la structure de ce type d'hiérarchie est basé sur la proximité des concepts.

e) La hiérarchie formelle (is-a)

C'est une hiérarchie dont la structure est déterminée par des relations de subsomption. Dans

Rapport-Gratuit.com

cette catégorie, si A est une super-classe de B, alors si un objet est une sous-classe de B, il est nécessairement le cas qu'il est ainsi une sous-classe de A. De même, si A est une superclasse de B, alors si un objet est une instance de B, alors il est nécessairement le cas qu'il est une instance de A. Cette ontologie ne peut inclure que des noms des classes.

f) La hiérarchie formelle (is-a) avec instances du domaine

Elle est similaire à la catégorie précédente mais en incluant des instances.

g) Le frame

C'est une ontologie incluant des classes avec leurs propriétés. Le fait d'inclure des propriétés dans la description d'une classe devient intéressant dans la mesure où ces propriétés peuvent être héritées par les sous classes de la taxonomie formelle « isa ».

h) Les ontologies avec restrictions de valeur

Ce sont des ontologies pouvant contenir des restrictions sur les valeurs des propriétés.

i) Les ontologies avec contraintes logiques

Ce sont les ontologies les plus expressives. Ces ontologies pouvant contenir des contraintes entre constituants (les relations) définies dans un langage logique. L'ontologiste peut exprimer les contraintes dans un langage d'ontologie expressif tel que OWL.

I .7.2. Typologie selon l'objet de conceptualisation

Gomez et al [14] proposent une classification selon le sujet de conceptualisation des ontologies. Cette classification est une extension des travaux de Mizoguchi [15], ceux de Van Heijst [16], et ceux de Guarino [17]:

a) Les ontologies de représentation de connaissances

Elles regroupent les primitives utilisées pour formaliser les connaissances sous un paradigme de représentation de connaissances.

L'exemple le plus expressif est l'ontologie de frame (Frame Ontologie) qui intègre les primitives de représentation des langages à base de frames:classes, instances, facettes, propriétés, relations, etc. [18]

b) Les ontologies générales ou communes ou génériques

Les connaissances modélisées dans ce type d'ontologie doivent être générales pour être réutilisées dans différents domaines. Elle comprend le vocabulaire relatif au temps, espace, unités, etc. [15]

c) Les ontologies de niveau supérieur ou de haut niveau

Les ontologies de niveau supérieur modélisent des concepts de haut niveau auxquels ces derniers doivent être reliés au sommet des ontologies de plus bas niveaux. Cependant, il existe plusieurs ontologies de haut niveau qui se diffèrent par le critère utilisé pour classifier les concepts généraux de la taxonomie.

d) Les ontologies du domaine

Elles sont réutilisables au sein d'un domaine donné, mais pas d'un domaine à un autre. Les connaissances représentées dans ce type d'ontologies sont spécifiques à un domaine particulier. Elles fournissent un vocabulaire d'un domaine spécifique au travers de, concepts et de relations qui modélisent les principales activités, les théories du domaine en question. Les concepts et les relations des ontologies de domaine sont souvent des spécialisations de concepts et des relations définis dans des ontologies de haut niveau. [16]

e) Les ontologies de tâches

Ce type d'ontologies est utilisé pour décrire un vocabulaire relatif à une tâche ou une activité générique (faire un diagnostic, planifier une activité …) en spécialisant certains termes des ontologies de haut niveau. Ces ontologies fournissent un ensemble de termes au moyen desquels on peut décrire, au niveau générique, comment résoudre un type de problème. [17]

f) Les ontologies d'application

Ce sont les ontologies les plus spécifiques. Contrairement à l'ontologie de domaine, l'ontologie d'une application donnée ne peut pas être réutilisée pour d'autres applications. Elle contient les connaissances requises pour une application particulière. Ce type d'ontologie décrit des concepts qui dépendent à la fois d'un domaine particulier et d'une tâche particulière. Par conséquent, elle spécialise souvent des ontologies de domaine et des ontologies de tâches pour une application donnée. [16]

I .7.3. Typologie selon le niveau de granularité

On peut distinguer les ontologies selon le niveau de détail utilisé lors de la conceptualisation de l'ontologie en fonction de l'objectif opérationnel envisagé pour l'ontologie, deux catégories peuvent être identifiées : granularité fine et granularité large. [19]

a) La granularité fine

Elle correspond à une ontologie très détaillée, possédant ainsi un vocabulaire plus riche capable d'assurer une description détaillée des concepts pertinents d'un domaine ou d'une tâche.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

b) La granularité large

Elle correspond à un vocabulaire moins détaillé. Les ontologies génériques possèdent une granularité large, compte tenu du fait que les notions sur lesquelles elles portent peuvent être raffinées par des notions plus spécifiques. [20]

I .7.4. Typologie selon le niveau de formalisation utilisé

Les ontologies peuvent être distinguées en fonction du degré de formalisme utilisé pour les exprimer. Uschold et Grüninger [21] proposent une classification contenant les quatre catégories: Hautement informelles; Semi-informelles; Semi-formelles et Rigoureusement formelles.

- Hautement informelle : Elle est exprimée en langue naturelle (sémantique ouverte).
- Semi-informelle : Exprimées dans un langage naturel structuré et limité, c'est-à- dire que des patrons ont été mis en œuvre.
- Semi-formelle : Elle est exprimée dans un langage artificiel défini formellement.
- Rigoureusement formelle : L'ontologie est exprimée dans un langage contenant une sémantique formelle, des théorèmes, et des preuves pour vérifier les propriétés telles que la validité et la complétude.

Dans l'étape suivante, nous nous intéressons d'abord au processus de développement d'ontologie, ensuite nous décrivons, analysons et comparons les différentes méthodes et méthodologies de développement des ontologies.

Dans la littérature, les mots méthodologie, méthode, technique, processus, activité sont bien souvent utilisés de façon aléatoire [22]. Pour clarifier l'utilisation de ces mots dans le domaine des ontologies, nous adoptons les définitions de l'IEEE (Institute of Electrical and Electronics Engineers). Ainsi les travaux sont présentés uniformément et indépendamment selon que chaque auteur utilise sa propre terminologie.

L'IEEE définit une méthodologie comme une série complète et intégrée de techniques ou de méthodes créant une théorie générale de systèmes de la façon dont un travail bien pensé doit être exécuté [23]. Les méthodologies sont largement utilisées en informatique et l'Ingénierie de Connaissance.

Une méthode est un ensemble de processus ou de procédures ordonnés utilisés dans la conception d'un produit ou dans l'exécution d'un service. [23]

Une technique est une procédure technique et managériale utilisée pour atteindre un objectif donné [23].

Selon la définition de l'IEEE, les méthodes et les techniques sont liées. En effet, les

méthodes et les techniques sont utilisées pour effectuer les tâches dans les différents processus composant une méthodologie.

Un processus est une fonction qui doit être exécutée dans le cycle de vie logiciel. Un processus est composé d'activités. [24]

Une activité est un ensemble de tâches constituant un processus [24].

Une tâche constitue la plus petite unité dans un travail. Une tâche est une partie de travail bien définie réalisée par une ou plusieurs personnes dans un projet. Les tâches liées sont habituellement groupées pour former des activités.

Pour synthétiser, nous pouvons dire qu'une méthodologie est composée de méthodes et de techniques. Les méthodes sont composées de processus et sont détaillées avec des techniques. Les processus sont composés d'activités. Et enfin les activités regroupent les tâches.

I .8. Méthodes et Méthodologies de Construction d'ontologie

I .8.1. Le processus de développement d'ontologie

Les activités exécutées lors de la construction des ontologies forment le processus de développement d'ontologie. Les ontologies sont construites par des équipes coopératives géographiquement distantes, et pour cette raison il est indispensable d'identifier ces activités. Le processus de développement d'ontologie est une proposition basée sur le standard IEEE pour le développement logiciel [25]. Il est recommandé d'effectuer les trois catégories des activités indiquées ci-dessous et d'orienter clairement le processus des constructions :

- Les activités de gestion incluant la programmation, le contrôle et la garantie de la qualité.
- Les activités orientées développement regroupant les activités de prédéveloppement, développement et post-développement.
- Les activités supports incluent une série d'activités exécutées en même temps que les activités orientées développement, sans lesquelles l'ontologie ne pourrait être construite. Ces activités incluent l'acquisition de connaissance, l'évaluation, l'intégration, la fusion, l'alignement, la documentation et la gestion de version.

Pendant le pré-développement, une étude de l'environnement est effectuée pour connaître l'environnement dans lequel l'ontologie sera utilisée, les applications dans lesquelles l'ontologie sera intégrée, etc. Aussi pendant le pré-développement, l'étude de faisabilité répond aux questions telles que : est-il possible ou est-il approprié de construire l'ontologie ?, etc.

Dans le développement, les activités à réaliser sont :

- La spécification ;

- La conceptualisation ;
- La formalisation ;
- L'implémentation.

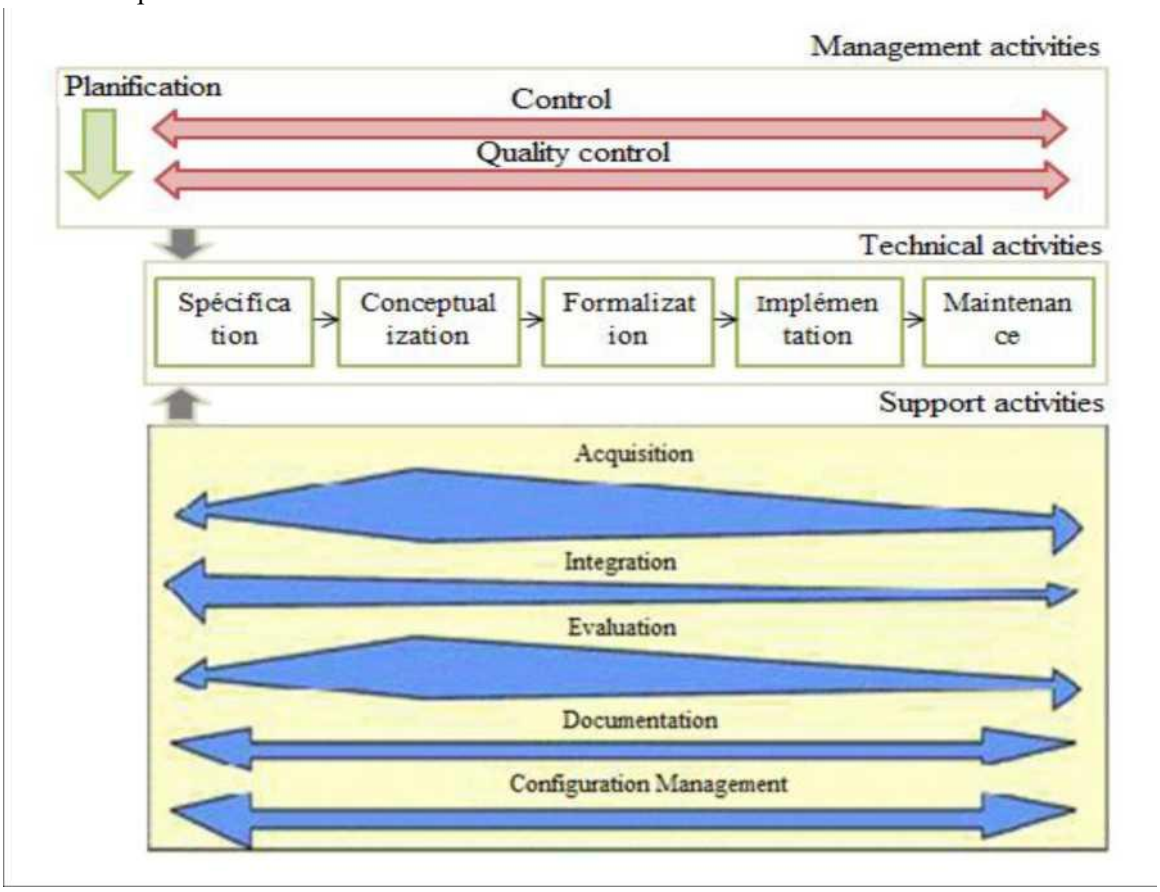

Figure I. 4 : Processus de développement d'ontologie de METHONTOLOGY. [25]

Pendant le post-développement, les activités de maintenance et de (ré) utilisation sont réalisées.

Le cycle de vie de l'ontologie identifie à quel moment les activités doivent être effectuées, c'est à dire, il identifie l'ensemble des étapes à travers lesquelles l'ontologie évolue, décrit quelles activités doivent être exécutées dans chaque étape et comment les étapes sont reliées.

I.8.2. Les méthodes et méthodologies de développement des ontologies

Une ontologie est toujours liée à une méthodologie de construction, à un outil de construction et avec un langage de représentation d'ontologie. A ce niveau, nous présentons les principales méthodologies et méthodes utilisées pour construire les ontologies à partir de zéro, par la réutilisation d'autres ontologies, par un processus de fusion ou par l'utilisation d'une approche d'étude d'ontologie. Ensuite nous nous intéressons en priorité aux méthodes et méthodologies utilisées dans le domaine de l'ingénierie de connaissance.

a) La méthodologie METHONTOLOGY

Cette méthodologie a été développée par le groupe d'Ontologie à l'Université Polytechnique de Madrid. METHONTOLOGY [25], elle prend ses racines dans les activités principales identifiées par le processus de développement logiciel et dans les méthodologies d'ingénierie de connaissance. Cette méthodologie inclue : l'identification du processus de développement, un cycle de vie basé sur des prototypes évolutifs, et les techniques pour effectuer chaque tâche dans les activités de gestion, de développement et de support.

ODE [26] et Web ODE [27] ont été construits pour donner un support technique à METHONTOLOGY. D'autres outils d'ontologie et de suites d'outils peuvent également être utilisés pour construire des ontologies avec cette méthodologie.

METHONTOLOGY a été proposée pour la construction d'ontologie par la FIPA (Foundation for Intelligent Physical Agents), qui favorise l'interopérabilité à travers les applications.

METHONTOLOGY inclut une méthode de réingénierie pour résoudre certains des problèmes liés à la construction d'une ontologie par la réutilisation d'une autre ontologie. [28]

b) La méthodologie On-To-Knowledge

On-to-Knowledge recommande un procédé itératif de développement, et comporte quatre phases principales : une phase de spécification de condition, une phase d'amélioration, une phase d'évaluation et une phase d'application et d'évolution. [28]

On-To-Knowledge propose l'acquisition des connaissances en spécialisant une ontologie générique. Elle propose de construire l'ontologie en tenant compte de la manière dont elle sera utilisée dans d'autres applications. Par conséquent, les ontologies développées avec cette méthodologie sont fortement dépendantes de l'application.

On-To-Knowledge recommande la suite d'outils OntoStudio comme support de développement des ontologies. Le principal auteur est GRÜNINGER [29]. Il indique clairement les relations entre un ensemble de techniques, méthodes et principes pour chacun de ses processus (par exemple, ordre recommandé, entrée/sorties). [30]

c) La méthode Ontology Development 101

OntologyDevelopment 101 [31] a été développée à l'Université de Stanford, elle cherche à construire des ontologies formelles par la reprise et l'adaptation des ontologies déjà existantes, et propose de suivre les démarches ci-après:

- Déterminer le domaine et la portée de l'ontologie ;
- Considérer la réutilisation des ontologies existantes ;
- Énumérer les termes les plus importants dans l'ontologie ;

- Définir les classes et hiérarchie des classes ;
- Définir les propriétés des classes ;
- Définir les facettes des attributs ;
- Construire les instances.

d) Les autres méthodes de développement d'ontologies

Ils existent d'autres méthodes de développement d'ontologies utilisées dans d'autres domaines. Certaines sont basées pour sur des études comparatives des méthodes de développement d'ontologie à partir de zéro. Dans ce groupe, nous pouvons citer entre autres :

- La méthode Cyc dans le domaine de la microélectronique. [72]
- Les méthodes d'Uschold [32], King [33] et de Gruninger [34], dont toutes les trois cherchent à modéliser le domaine de l'entreprise. Elles sont moins détaillées et n'ont été testées que dans le domaine des affaires (business).
- La méthode Kactus [35], proposée dans le domaine de l'intelligence artificielle ;
- La méthode Sensus, proposée par l'équipe de Swartout [36], est utilisée pour construire le squelette d'une ontologie de domaine à partir d'une grande ontologie.

D'autres méthodes, qui ne s'occupent que d'une activité précise du processus de développement existent. On peut les classer dans les catégories suivantes:

- Les méthodes d'étude des ontologies ou «ontologies learning». Cette catégorie de méthodes s'intéresse à l'activité de l'acquisition de connaissance. L'acquisition de la connaissance pour la construction des ontologies exige toujours beaucoup de temps et de ressources, malgré que l'un des principaux objectifs des ontologies soit de focaliser et réduire le niveau de l'acquisition de la connaissance

Ces méthodes sont utilisées pour créer une nouvelle ontologie à partir de zéro, enrichir une ontologie existante avec de nouveaux termes, et acquérir la connaissance pour des tâches.

- Les méthodes d'alignement et de fusion des ontologies. Les ontologies ont pour but de capturer la connaissance consensuelle d'un domaine donné d'une manière générique et formelle pour être réutilisée et partagée à travers des applications et par des groupes de personnes. De cette définition nous pourrions incorrectement déduire qu'il y a seulement une ontologie pour un domaine. En fait, nous pouvons trouver dans la littérature plusieurs ontologies qui modélisent, de différentes manières, le même type de connaissance ou de domaine. Noy et Musen [37] distinguent deux approches pour unifier les terminologies des ontologies : alignement d'ontologie et fusionnement d'ontologie. Les méthodes d'alignement d'ontologie établissent différents types de correspondances entre les

ontologies, par conséquent cette option préserve les ontologies originales.

Plusieurs méthodes d'alignement d'ontologie ont été proposées, comme celle utilisée par Anchor PROMPT [38]. Cependant, les méthodes de fusion d'ontologies proposent de produire une ontologie unique à partir des ontologies originales. Le processus de fusionnement exige habituellement d'établir des correspondances entre les ontologies afin de les fusionner. Etant donné l'état actuel des choses et dans le contexte du Web Sémantique, il est plus approprié d'établir des correspondances ontologiques entre les ontologies existantes sur le même sujet que de prétendre construire un modèle unifié de la connaissance pour un tel sujet à partir de zéro. Dans cet ordre d'idées, nous pouvons citer la méthodologie OIGNONS [39], les méthodes FCA-Merge [40] et PROMPT [31] pour fusionner des ontologies.

- Les méthodes d'évaluation des ontologies sont des ontologies qui ne peuvent être réutilisées par d'autres ontologies ou être utilisées par des applications sans une évaluation préalable de leur contenu d'un point de vue technique. Comme directives pour évaluer des ontologies, les premiers travaux ont été proposés par l'équipe de Gomez-Perez [28]; et ont continué sur l'évaluation de la connaissance taxonomique. L'équipe de Guarino a développé OntoClean [4], qui est une méthode pour analyser et nettoyer la taxonomie d'une ontologie existante au moyen d'un ensemble de principes basés sur la philosophie.

I .9. Domaine d'application des ontologies

I .9.1. Système d'information

L'intégration d'une ontologie dans un système d'information vise à réduire, voire à éliminer, la confusion conceptuelle et trilogique à des points clefs du système, et a tendre vers une compréhension partagée pour améliorer la communication, le partage, l'interopérabilité et le degré de réutilisation possible, ce qui permet de déclarer formellement un certain nombre de connaissances utilisées pour caractériser les information gérées par le système, et de se baser sur ces caractérisations et la formalisation de leur signification pour automatiser des tâches de traitement de l'information.

L'ontologie se retrouve maintenant dans une large famille de système d'information.

Elle est utilisée pour :

- Décrire et traiter des ressources multimédia ;
- Assurer l'interopérabilité d'application en réseau ;

- Construire des solutions multilingues et interculturelles ;
- Permettre l'intégration des ressources hétérogènes d'information ;
- Vérifier la cohérence de modèle ;
- Permettre les raisonnements temporels et spatial ;
- Faire des approximations logiques; etc.

Ces utilisations des ontologies se retrouvent dans de nombreux domaines d'application tel que:

- Intégration d'information géographique ;
- Gestion de ressources humaines ;
- Aide à l'analyse en biologie ;
- Commerce électronique ;
- Enseignement assisté par ordinateur ;
- Bibliothèque numérique ;
- Recherche d'information. Etc...

I .9.2. Web sémantique

Un courant particulièrement promoteur pour l'expansion des systèmes à base d'ontologies est celui du web sémantique. Il s'agit d'une extension du web actuel, dans laquelle l'information se voit associée à un sens bien définit, améliorant la capacité des logiciels à traiter l'information disponible sur le web. L'annotation des ressources d'informations du web repose sur des ontologies, elles sont aussi disponibles et échangées sur le web.

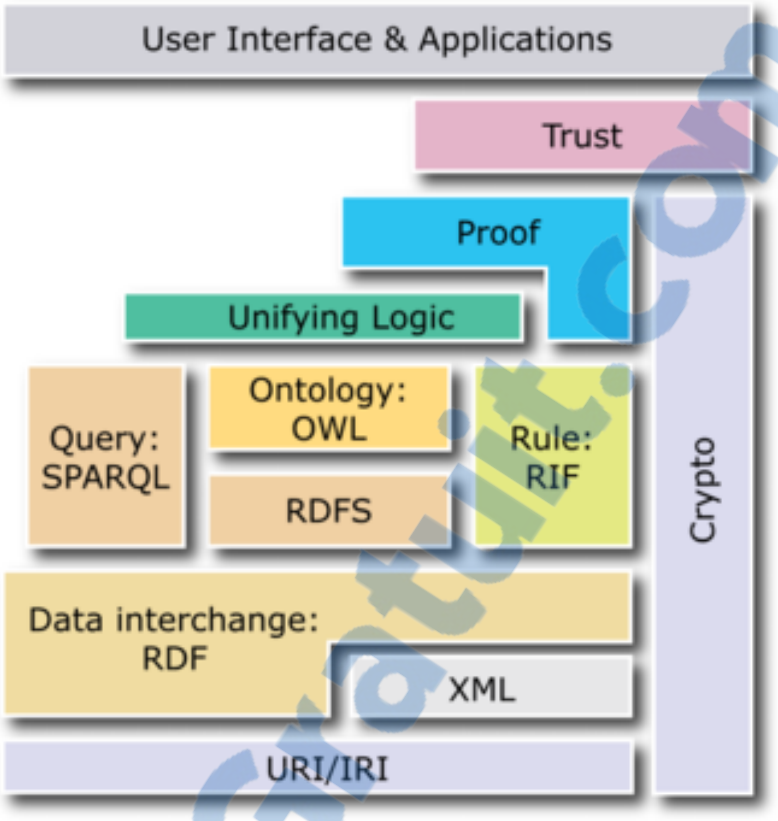

Figure I. 5 : Pyramide du web sémantique. [41]

Dans ce paragraphe nous donnons la structure de la pyramide du web sémantique.

La base de la pyramide regroupe les technologies classiques, comme :

Les URI (Universal Resource Identifier) permettant, d'identifier de façon unique des ressources. Ces ressources ne sont pas nécessairement disponibles sur le web.

Le langage XML (extensible Markup Langage) qui permet d'écrire des documents semi-structurés. La plupart des langages normalisés par le W3C dans le cadre du Web sémantique sont des dialectes XML. [41]

La couche du milieu contient des technologies développées spécifiquement dans le cadre du web sémantique. Tout d'abord, le langage RDF (Resource Description Framework) est un format d'échange XML utilisé pour la représentation des connaissances sur les ressources du Web. C'est en quelque sorte le langage de base du Web sémantique. RDF permet de voir le Web comme un ensemble de ressources reliées par les liens étiquetés sémantiquement. Il permet de décrire des entités avec leurs propriétés, chacune associée à sa valeur, à partir d'un modèle simple de triplet (ressource, propriété, valeur). Au-dessus de RDF se trouve RDF Schéma (RDFS) qui s'est ajouté rapidement pour offrir un niveau supérieur de structuration.

Rapport-Gratuit.com

La pyramide du W3C inclut aussi des couches supérieures où se trouvent des langages plus complets et plus complexes que RDFS. Parmi ces langages, on peut citer:

- OWL (Web OntologyLanguage) qui est une famille de langages de représentation des connaissances pour définir des ontologies. Basé sur les logiques de description, OWL enrichit les possibilités de description et de raisonnement offertes par RDF Schema.
- Le langage SPARQL (SPARQL Protocol and RDF QueryLanguage) est un langage de requêtes pour RDF. Ce langage permet d'extraire des informations dans les documents RDF.
    - Le langage RIF (Rule Interchange Format), est une recommandation du W3C. RIF fait partie de la feuille de route du web sémantique, avec principalement SPARQL, RDF et OWLL. Bien que prévu initialement par beaucoup comme une "couche de règles logiques" pour le web sémantique, en réalité, la conception du RIF est basée sur l'observation qu'il existe de nombreux "langages de règles" et qu'il est nécessaire de s'échanger des règles entre eux. [73]
    - Les parties « UnifingLogic » et « Proof » forment la couche logique du Web sémantique, destinée au support des règles et à l'unification dans les ontologies. Les composants « Cryptography » et « Trust » (confiance) ont pour rôle la gestion de la sécurité des données sur le Web sémantique.

I .10. Conclusion

Nous avons présenté les ontologies avec leurs caractéristiques, leurs constructions et leurs méthodes d'application, ainsi que les domaines d'application de ces ontologies. Dans ce qui a été présenté dans ce chapitre, il ressort que la notion d'ontologie constitue l'une des approches les plus efficaces pour représenter et analyser et traiter des connaissances.

La démarche d'alignement étant basée sur trois processus :

Le processus d'extraction, qui présente l'étape de découverte de la sémantique correspond à la modélisation de la sémantique, le processus de représentation et enfin le processus d'alignement proprement dit. **Le processus d'extraction de la sémantique** est effectué sur la base d'ontologies. En se basant sur des critères de similitudes le processus de représentation d'ontologie nous permet de quantifier le degré de similarité des concepts des ontologies et enfin le **processus d'alignement** nous permettra d'aligner les ontologies concernées pour une analyse et traitement optimum des connaissances.

Chapitre II

Les langages et les outils d'ontologie

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

II .1. Introduction

Le traitement et l'analyse de l'information a été considérablement amélioré par l'outil informatique. La recherche d'information avec ses variations sémantiques, la gestion des connaissances, les systèmes coopératifs, le Web sémantique etc.… ont été considérablement progressés avec le développement des ontologies. L'efficacité d'un développement d'ontologie dépend de façon très particulière du langage d'implémentation utilisé. Le choix d'un langage d'ontologie n'est pas basé sur la représentation de la connaissance et sur les mécanismes d'inférence exigés par l'application qui utilise l'ontologie, mais il est basé sur les préférences du développeur.

Le besoin, en termes d'expressivité et de raisonnement doit être parfaitement connu avant le codage de l'ontologie pour définir les langages qui satisfont à de telles exigences.

Dans ce chapitre, nous faisons une description des langages utilisés pour le développement des ontologies et des outils pour leurs constructions.

II .2. Les langages de description des ontologies

Nous pouvons classer les langages de description d'ontologies existants en deux catégories, à savoir traditionnels et liés à internet.

II .2.1. Les langages d'ontologies traditionnels

Cette catégorie représente un ensemble de langages d'ontologie basés sur l'Intelligence Artificielle, créés la plupart dans les années 1990. Nous pouvons citer entre autres :

- CycL basé sur les frames et la logique de premier ordre, et utilisé pour la construction de l'ontologie Cyc.
- KIF (Knowledge Interchange Format), désigné comme un format d'échange de la connaissance, et basé sur la logique de premier ordre.
- Ontolingua basé sur KIF, il est aujourd'hui le langage d'ontologie supporté par Ontolingua Server, et il est basé sur les frames et la logique de premier ordre.
- LOOM basé sur les logiques de description et les règles de production, n'était pas prévu pour implémenter des ontologies mais pour les bases de connaissances générales. Ce langage fournit des dispositifs de classification automatique des concepts.
- OCML (OperationalConceptual Modeling Language), une sorte d'Ontolingua opérationnel, parce que la plupart des définitions qui peuvent être exprimées dans OCML sont similaires aux définitions correspondantes dans Ontolingua. Ce langage a été construit pour développer des ontologies exécutables et des modèles dans les méthodes de résolution de problème.

- Flogic(Frame Logic), développé comme un langage qui combine les frames et la logique de premier ordre.
- OKBC (Open Knowledge Base Connectivity) un protocole né dans le cadre d'un programme de recherche de DARPA. Ce protocole permet d'accéder aux bases de connaissance stockées dans différents systèmes de représentation de connaissance, qui peuvent être basés sur différents paradigmes de représentation de connaissance. Parmi les langages mentionnés ci-dessus seuls Ontolingua, LOOM et CycL sont conformes à OKBC.

La disposition globale de ces langages est montrée sur la figure II. 1

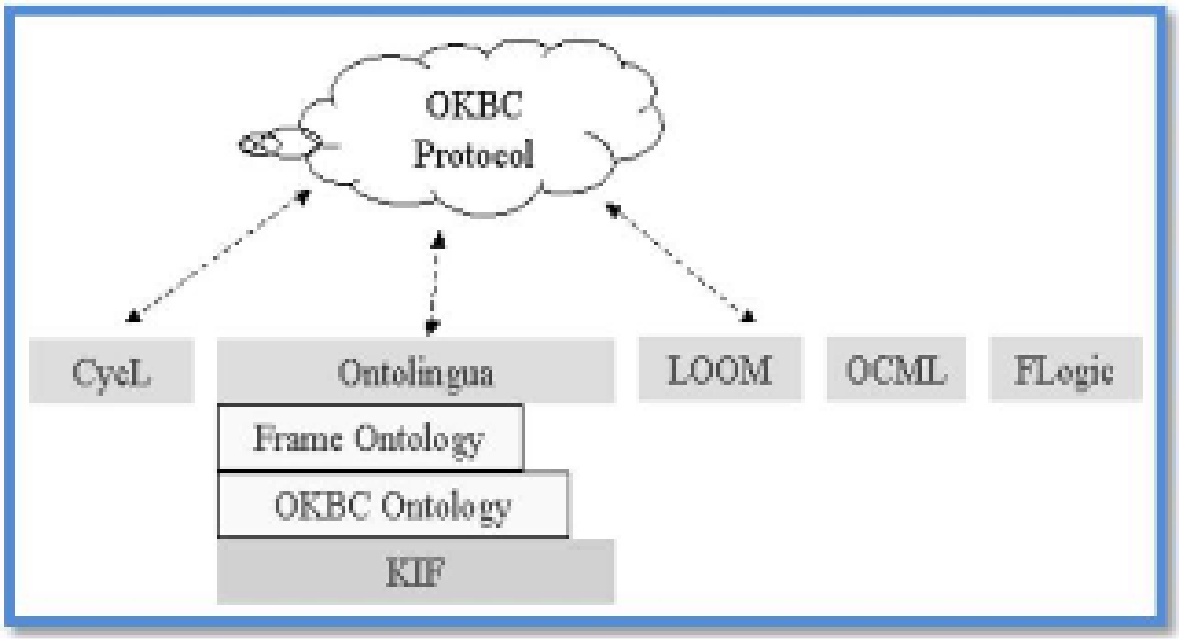

Figure II. 1: Les langages d'ontologies traditionnels [28]

II .*2.2.* Les langages d'ontologie web standard ou basés sur XML

La création de langages d'ontologie pour exploiter les caractéristiques du Web a été favorisée par l'utilisation massive de l'Internet. Ces langages, appelés des langages web standard ont une syntaxe basée sur les langages existants tels que HTML et XML, dont les buts ne sont pas le développement d'ontologie mais la présentation et l'échange des données respectivement.

Les relations entre ces langages sont montrées dans la figure II.2.

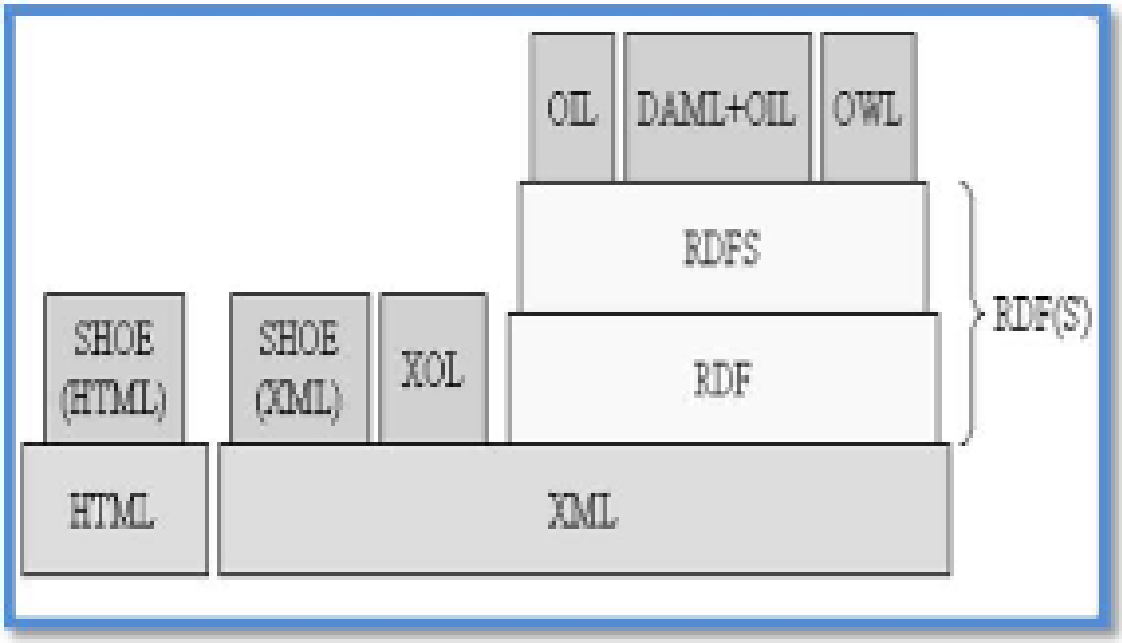

Figure II.2 : Les langages d'ontologies basés sur xml [28]

Dans cette catégorie de langages, nous pouvons citer :

- SHOE (Simple HTML OntologyExtension), basé sur le web, il combine les frames et les règles. Il a été construit comme une extension de HTML.
- XOL (XML-basedOntology exchange Language), développé comme une transformation en XML d'un petit sou ensemble des primitives du protocole OKCB.
- RDF (Resource Description Framework), développé par le W3C comme un langage basé sur les réseaux sémantiques pour décrire les ressources web.
- RDF Schéma, construit par le W3C comme une extension de RDF avec des primitives basées sur les frames.

La combinaison de RDF et de RDF Schéma est connue sous le nom de RDF(S).

Les langages ci-dessus ont établi les bases du Web Sémantique. Et c'est dans ce contexte que trois autres langages ont été développés comme des extensions de RDF(S). Il s'agit de :

- OIL (Ontology Interchange Language and Ontology Inference Layer).
- DAML+OIL (DARPA Agent Markup Language).
- OWL (OntologyWeb Language), créé par le W3C comme le nouvel langage pour Web sémantique.

II .3. Les outils de construction d'ontologie

On distingue deux types d'outils pour la construction des ontologies : les outils de construction d'ontologie dépendants du formalisme de représentation et les outils de construction

d'ontologie indépendants du formalisme de représentation.

II .3.1. Les outils dépendants du formalisme de représentation

a. Ontolingua [74]

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. Le langage Ontolingua utilise des classes, des relations, des fonctions, des objets (instances) et des axiomes pour décrire une ontologie. Une relation (une classe) peut contenir des propriétés nécessaires (contraintes) ou nécessaires et suffisantes qui définissent la relation. [42] [43]

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 : on redéfinit une définition importée depuis n'importe quelle ontologie.

b. OntoSaurus [75]

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'usager d'éditer l'ontologie. Il utilise LOOM comme langage de représentation des connaissances. On peut représenter les concepts, la taxonomie des concepts, les relations entre les concepts, les fonctions, les axiomes et les instances.

c. WebOnto [76]

WebOnton est une application Web pour naviguer et développer de façon collaborative les ontologies. WebOnto supporte la navigation collaborative, la création et l'édition d'ontologies sur le Web. Les ontologies WebOnto sont implémentées dans le langage OCML, qui est une combinaison des frames et de la logique de premier ordre et qui permet de représenter les concepts, la taxonomie des concepts, les relations, les fonctions, les axiomes et les instances.

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 permettant 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.

d. OilEd [44]

OilEd (Oil Editor) est un éditeur d'ontologies utilisant le formalisme OIL. Il est essentiellement dédié à la construction de petites ontologies dont on peut 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. [44]

II .3.2. Les outils indépendants de formalisme de représentation

a. Protégé2000 [77]

Protégé2000 est une interface modulaire permettant l'édition, la visualisation, le contrôle 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 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. [45]

b. ODE et WebOde [46] [47]

L'outil ODE (OntologyDesign 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. [46] [47]

c. OntoEdit [48]

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 généricité 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. [48]

II .4. Conclusion

Dans ce chapitre, nous avons présenté les différents langages d'implémentation d'ontologies, en insistant sur le fait que le choix du langage d'implémentation utilisé joue un rôle important dans la construction de l'application donnée. Nous avons aussi présenté une liste d'outils

Chapitre III

Etat
De l'art sur les ontologie

Dans cette partie, nous présentons un début état de l'art sur les ontologies. D'abord nous donnons les différentes distinctions entre les ontologies et les terminologies, les différents langages tels RDF, RDF(S), OWL ... Ensuite, nous décrivons les méthodes et méthodologies de construction d'ontologie. Puis nous nous intéressons aux maintenances et évolution d'ontologies. Enfin nous présentons les différents exemples d'ontologies et de classifications dans le domaine médical.

III.2. Terminologies

Les ontologies terminologiques sont définies comme étant des ontologies dont les catégories ne sont pas spécifiées par des axiomes et des définitions. Un exemple est Word Net [49] dont les catégories sont partiellement spécifiées par des relations comme l'hyperonymie, qui détermine les positions relatives des concepts, sans pour autant la définir de façon complète. La plupart des organisations établissent des conventions de nommage, des classifications et des standards. Ces éléments sont un début de formalisation.

III.3. Ontologies Vs Terminologies

Une distinction entre terminologie et ontologie est apportée par [50] C. Roche qui stipule que « ...parler de terminologie c'est parler des mots et du sens des mots... L'étude du sens d'un mot relève de la linguistique... ». Dans une terminologie on s'intéresse aux mots et aux relations entre eux ; la relation structurante de base est la relation d'hyperonymie (Rapport d'inclusion entre des unités lexicales, considéré comme orienté du plus général au plus spécifique) et son inverse l'hyponymie(est la relation sémantique d'un lexème à un autre selon laquelle l'extension du premier est incluse dans l'extension du second), tandis que dans une ontologie, on s'intéresse à la notion de concept et aux relations entre eux. [50]

Cependant, ontologies et terminologies servent au même objectif ; fournir à une communauté d'utilisateurs une conceptualisation partagée d'une partie spécifique du monde, dans le but de faciliter une communication efficiente d'une connaissance complexe. Gamper fournit quatre critères pour distinguer les deux notions [51] :

- Le cadre formel de leur définition : le sens des termes est défini dans un texte libre en langue naturelle. L'interprétation correcte du sens véhiculé dépend de l'utilisateur. Les ontologies spécifient explicitement la connaissanceConceptuelle en utilisant un langage formel avec une sémantique claire, qui permet une interprétation non ambiguë des

Rapport-Gratuit.com

termes ;

- Le support computationnel : La plupart des terminologies utilisées actuellement fournissent très peu de sémantique pour une représentation explicite aussi bien de la connaissance que pour la maintenance des données. Pour les ontologies, quant à elles et grâce à l'utilisation de langages de représentation formels comme les Logiques de Description (DLs) [52], il est possible de vérifier leur consistance, d'inférer de nouvelles connaissances, etc. ;
- les utilisateurs : les terminologies sont utilisées pour véhiculer l'information entre les humains, tandis que les ontologies sont principalement développées pour le partage de connaissances entre l'être et la machine.
- l'usage de la langue naturelle : la terminologie s'occupe du transfert de la connaissance comme une activité linguistique, c'est-à-dire l'exploration de la langue naturelle afin d'identifier tous les termes utilisés par les humains pour faire référence aux concepts sous-jacents.

La question de savoir si une ontologie donnée doit être représentée de façon formelle ou informelle est directement liée à l'objectif de cette ontologie [53]. Si le but est d'assurer la communication entre des personnes, une description en langue naturelle peut suffire. Par contre, si elle doit assurer le rôle d'un format d'échange dans un environnement d'interopérabilité entre programmes, il est nécessaire qu'elle soit formelle afin que la communication soit cohérente et que la vérification automatique de cette cohérence soit possible.

III.4. Représentation des ontologies

Plusieurs langages, dont la syntaxe est basée sur le langage XML, ont été conçus pour une utilisation des ontologies dans le cadre du web sémantique. Parmi eux, les plus importants sont RDF et RDF(S) [54], qui est la recommandation du W3C pour représenter les métadonnées, OIL, DAML+OIL, OWL qui est la recommandation du W3C pour représenter des ontologies. Nous allons dans la section suivante faire une revue des langages RDF et RDF(S) et OWL, utilisés dans la cadre du Web Sémantique.

III .4.1. RDF et RDF(S)

RDF (Resource Description Framework) est né d'une recommandation du W3C pour décrire des ressources. Il représente un graphe pour la description de métadonnées en permettant leur traitement automatisé. A l'origine, il a été défini pour décrire des ressources du Web telles que les pages Web ; aujourd'hui il peut décrire toute chose ayant une identité.

RDF décrit les ressources en exprimant des propriétés et en leur attribuant des valeurs. Il utilise pour cela le vocabulaire défini par RDF Schémanoté RDF(S).

RDF(S) fournit un ensemble de primitives simples, mais puissantes, pour la structuration de la connaissance d'un domaine en classes et sous-classes, propriétés et sous-propriétés avec la possibilité de restreindre leur domaine d'origine (rdf : Domain) et leur domaine d'arrivée (rdf : range).

L'élément de base d'un document RDF est la ressource, correspondant à la représentation conceptuelle d'une entité. Une ressource est identifiée par son URI (Uniform Resource Identifier), de la forme http : //uri/du/document#courant.

RDF repose sur un modèle de triplet (sujet, propriété, objet) ou (sujet, verbe, complément) [55].

- Le sujet est la ressource à décrire;
- L'obj et est une donnée ou une autre ressource;
- Chaque triplet est représentable par un graphe propriété

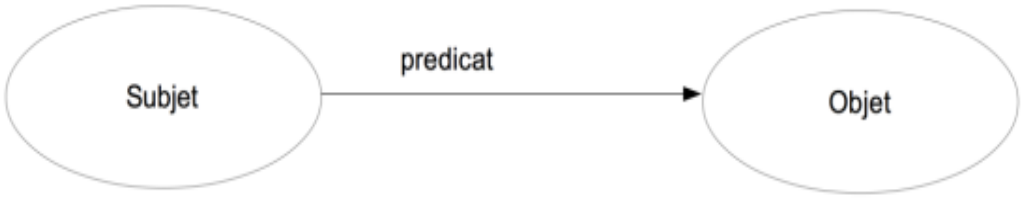

Figure 1Une triple projection de sujet, prédicat et objet

III .4.2. OWL

RDF/RDF(S) ne permet pas d'exprimer certaines notions que l'on voudrait décrire avec les ontologies. Les chercheurs du W3C ont proposé le langage d'ontologie du Web (OWL), basé sur le langage DAML+OIL [56]. La sémantique du langage (ou plus exactement celle de OWL DL et OWL Lite) peut être définie via une transformation en Logiques de Description (DL), dont ce langage est inspiré. On peut utiliser les outils de raisonnement développés pour les Logiques de Description afin d'effectuer des inférences sur OWL.

Ces chercheurs ont défini trois différents sous-langages de OWL, et ceci dans le but de répondre aux critères exigés pour un langage de représentation d'ontologies :

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

a. OWL Full

OWL Full est le langage complet. Il utilise tous les éléments disponibles en OWL. Il permet aussi de combiner de façon quelconque ces éléments avec RDF et RDF Schema. OWL Full est totalement compatible avec RDF ; ainsi tout document RDF valide est aussi un document OWL Full valide. Son inconvénient principal vient de son pouvoir d'expression élevé, qui le rend indécidable.

b. OWL DL

OWL DL (OWL Description Logic) qui offre une plus grande expressivité que OWL Lite et il est décidable. Il est ainsi appelé à cause de son origine, le formalisme de représentation des connaissances appelé Logique de Description.

Une ontologie consiste donc en un ensemble d'axiomes qui expriment les relations par exemple la relation de subsumption entre les classes ou les propriétés. Chaque élément de l'ontologie appartient à une seule catégorie, c'est-à-dire qu'un même objet ne peut être défini à la fois comme une classe et comme une propriété. OWL DL a l'avantage de fournir un support pour les inférences, mais la compatibilité totale avec RDF et RDF(S) est perdue.

c. OWL Lite

OWL Lite permet de représenter des classifications sous forme hiérarchique et d'exprimer des contraintes simples (contraintes de cardinalité de type 0 ou 1). La disjonction de classes, la définition de classes à partir d'une union de classes, et bienautres possibilités offertes par OWL DL, qui ne sont pas autorisées. Ce langage est particulièrement indiqué pour représenter des taxonomies et autres thésaurus.
Il y a une stricte compatibilité ascendante de ces trois langages :

- Toute ontologie OWL Lite valide est une ontologie OWL DL valide ;
- Toute ontologie OWL DL valide est une ontologie OWL Full valide ;
- Toute conclusion d'une ontologie OWL Lite est une conclusion valide de OWL DL ;
- Toute conclusion OWL DL est une conclusion valide de OWL Full.

III.5. Construction et édition d'ontologies

La construction d'une ontologie n'est pas une activité aisée, d'autant plus qu'il n'existe pas aujourd'hui une méthodologie communément admise, comme c'est le cas dans le domaine des bases de données par exemple, avec la démarche Entité/Association. Le processus de construction intègre souvent un expert du domaine. Une fois construite, l'ontologie doit être

claire, cohérente, compréhensible, facile à utiliser et extensible [4].

Nous pouvons classer les travaux sur la construction des ontologies comme suit [57]:

- les méthodes et méthodologies pour la construction d'ontologies en partant de zéro ;
- les méthodes pour la réingénierie d'ontologies ;
- les méthodes de construction coopérative d'ontologies.

Corcho et ses collègues [58] et Murshed et Ali [59] ont fait une description très détaillée des méthodes et outils sur la construction d'ontologies. Corcho et ses collègues ont fait une revue des principaux outils et langages ainsi que des méthodologies pour construire des ontologies, tandis que Murshed et Ali ont proposé des critères d'évaluation et des techniques de classement permettant aux développeurs d'ontologies de choisir le bon outil de construction.

III.5.1. Les méthodes et méthodologies pour la construction d'ontologies en partant de zéro

Dans cette catégorie, la première a été la méthodologie CYC [60], élaborée dans le cadre du développement de l'ontologie CYC. Cette méthodologie propose deux étapes de construction :

- l'extraction manuelle de la connaissance commune implicite dans les différentes sources ;
- l'utilisation des techniques de Traitement Automatique de la Langue Naturelle (TALN) et d'acquisition de connaissances pour générer de nouvelles connaissances à partir de celles acquises à l'étape précédente.

III .5.2. Les méthodes pour la réingénierie d'ontologies

Reconstruire et lier un modèle conceptuel d'une ontologie déjà implémentée à une autre en cours d'implémentation représente le processus de réingénierie d'ontologies, [61] [62]. En adaptant une technique de réingénierie de schémas à une ontologie de domaine Gomez-Perez [28], donne une méthode représentative de ce processus.

III .5.3. Les méthodes pour la construction coopérative d'ontologies

Les ontologies sont souvent utilisées par plusieurs praticiens, elles doivent donc faire l'objet d'un consensus et être acceptée par sa communauté d'utilisateurs. Leur construction est une construction coopérative. Les méthodes de cette catégorie adoptent donc une approche collaborative de construction incluant l'intervention de personnes localisées à des endroits différents. Un certain nombre de problèmes que pose la construction d'ontologies dans un

contexte distribué : la gestion de l'interaction et la communication entre les différentes personnes ; le contrôle de l'accès aux données ; la reconnaissance de l'attribution des droits d'auteurs ; la détection et la correction d'erreurs ont été identifié par Euzenat [63]. Comme exemple de ces méthodes citons celle de Farquhar et ses collègues [64], utilisée pour la construction d'ontologies pour l'intégration de données et la méthode élaborée dans le cadre de l'initiative KA [65], utilisée pour la construction d'ontologies pour l'annotation sémantique de documents.

III.6.Enrichissement, maintenance et évolution d'ontologies

Une ontologie dans son cycle de vie est appelée à évoluer afin de prendre en compte les spécificités des nouveaux besoins qui peuvent apparaître. Ce changement peut être de natures très différentes ; concernant la structure conceptuelle, l'ajout de nouveaux concepts, de nouvelles définitions, de nouveaux termes, etc. L'enrichissement d'une ontologie par de nouveaux termes pour désigner les concepts est particulièrement souhaité dans le cadre de l'utilisation d'ontologies pour la recherche d'information [66].

III.7.Exemples d'ontologies et de classifications dans le domaine médical

Le foisonnement des classifications médicales illustre la nécessité pour une communauté scientifique de disposer d'un vocabulaire de consensus. La multiplicité de ces classifications, particulièrement notable dans le domaine médical, illustre aussi le fait que ces besoins se sont manifestés dans différentes communautés, qui y ont répondu d'une manière opportuniste, en fonction de leurs moyens et de leurs besoins. D’où la présence de plusieurs classifications, répondant chacune à une catégorie de besoins, mais présentant des structurations différentes, bien que les concepts qu'elles définissent présentent un taux de recouvrement important. [66]

Pour tenter de remédier à cette situation, la NLM (National Library of Medicine) a pris en 1986 l'initiative du projet UMLS (UnifiedmedicalLanguageSystem) [67].

L'objectif d'UMLS est d'améliorer la capacité des programmes informatiques à comprendre la signification des requêtes des utilisateurs dans le domaine médical et d'utiliser cette compréhension pour la recherche et l'intégration d'informations compréhensibles par les machines.

Quelques autres initiatives sont significatives : le projet GALEN (GeneralisedArchitecture for Languages, Encyclopaediasand Nomenclatures in Medicine) pour le développement de systèmes multilingues de terminologies médicales, le projet Digital Anatomiste[68] pour la

conceptualisation des objets physiques et spatiaux qui constituent le corps humain à une échelle macroscopique [68], le thésaurus MeSH[78], utilisé pour l'indexation de ressources médicales en ligne (MEDLINE)[79], et SNOMED CT18[80], un vocabulaire contrôlé utilisé dans le domaine clinique.

III .7.1. UMLS [81]

UMLS est un recueil de plusieurs vocabulaires dans le domaine médical, il fournit une structure de cartographie de vocabulaires et permet la traduction entre des différents systèmes de terminologies.

Le Cœur d'UMLS est constitué d'un Meta thésaurus qui comporte plus de 1.300.000 concepts dans sa version 2006 et qui croît régulièrement de près de 200.000 concepts par an. C'est la partie la plus évolutive d'UMLS, et le rythme de sa croissance pose la question de sa maîtrise et de sa cohérence.

Chaque concept UMLS a un identifiant unique, le CUI (Concept Unique Identifier). A chaque concept est associé un ensemble de termes dans différents lexiques. Chaque CUI a dans chaque langue un terme préféré unique appeler SUI (String Unique Identifier). Chaque SUI est lié à un ou plusieurs termes selon ses différentes variations lexicales, qui sont les LUI (Lexique Unique Identifier). Les SUI dans les différentes langues sont nécessaires pour la communication, mais le vrai identifiant du concept est son CUI.

Même si UMLS constitue aujourd'hui une source terminologique très utilisée, certaines imperfections sont à relever. Gangemi et ses collègues [69] notent certains problèmes de cycles, qui peuvent être causés par le recouvrement partiel de certains concepts.[69]

III .7.2. SNOMED CT [80]

SNOMED-CT (SystematizedNomenclature of Medicine, ClinicalTerms) combine la SNOMED Reference Terminology (SNOMED RT) et la Version 3 de la United Kingdom'sClinicalTerms (anciennement connue sous le nom de Read Codes). Son usage se retrouve dans plusieurs contextes : l'indexation de documents cliniques, l'aide à la décision clinique, l'indexation d'images, etc. La terminologie est organisée en une hiérarchie de 18 catégories de premier niveau : les procédures, les entités observables, les structures du corps humain, les événements, etc. Ces entités, dites majeures, sont regroupées autour d'une racine appelée Top. Dans sa version de Juillet 2006, SNOMED CT totalise plus de 300.000 concepts et 770.000 descriptions en anglais.

III .7.3. MeSH [78]

Les thésaurus sont des outils incontournables dans la construction des ontologies ; dans le domaine médical la National Library of Medicine(NLM) a réalisé le MeSH (MedicalSubjectHeadings). Il est utilisé pour l'indexation et la recherche d'informations médicales. La première version est apparue en 1954 sous le nom de SubjectHeadingAuthority List. Il sert de vocabulaire contrôlé pour l'indexation des ressources de la base de données bibliographique MEDLINE20. Il est aussi utilisé par les portails d'indexation et de catalogage de ressources médicales.

Le nombre sans cesse croissant de ressources médicales à gérer par les documentalistes médicaux, a contraint la NLM à lancer le projet MEDLARS pour l'automatisation de l'indexation et la recherche de ressources médicales. Depuis, le MeSH n'a cessé d'évoluer.

Les descripteurs MeSH sont organisés en 16 catégories, qui sont elles aussi subdivisée en sous-catégories. A l'intérieur de chaque catégorie, les descripteurs sont structurés hiérarchiquement, du plus général au plus spécifique, avec un niveau de profondeur maximum de 11. Les catégories représentent les termes anatomiques, les maladies, etc.

III .7.4. Galen

GALEN est une initiative européenne pour le développement de systèmes multilingues de terminologies médicales. Galen est basée sur un modèle sémantique puissant de terminologie clinique, le Galen Coding Reference (CORE), et utilise un langage de représentation appelé GRAIL (GALEN Représentation And IntegrationLanguage) dont le noyau est basé sur les logiques de description. Ce modèle consiste en une hiérarchie de subsomption d'entités élémentaires (concepts primitifs des DL) et un ensemble de déclarations liant ces entités. Le modèle terminologique permet une définition explicite des concepts dans le domaine médical.

Comme toute ontologies, une ontologie Galen est composée d'une hiérarchie de catégories élémentaires, une hiérarchie de liens sémantiques appelés attributs (aussi désignés par rôles ou relations) pour établir des relations entre les catégories. Les concepts peuvent êtres simples ou composés. Les concepts composés sont une combinaison de concepts simples. Par ailleurs, on peut définir un ensemble d'axiomes sur les concepts, ainsi que des contraintes.

III .7.5. Digital Anatomist:Foundational Model of Anatomy (FMA)

Le projet Digital Anatomista pour objectif la conceptualisation des objets physiques et spatiaux qui constituent le corps humain à une échelle macroscopique. Une ontologie

anatomique du corps humain a été construite en définissant les classes et les sous- classes des entités anatomiques physiques selon leurs relations spatiales.

III.8. Conclusion

Nous avons essayé de donner un aperçu des notions liées aux ontologies. Cette aperçu, bien que n'étant pas exhaustife car ce domaine est assez vaste aujourd'hui, Nous avons décrit quelques langages de représentation des ontologies utilisés dans le cadre du Web sémantique. RDF et RDF(S) ont été conçus suivant le principe de décentralisation du Web en permettant de représenter, d'intégrer et de faire évoluer des vocabulaires distribués. Nous avons vu cependant que RDF(S) ne supportait pas certaines notions à propos des ontologies, comme l'expression de l'équivalence entre concepts, les contraintes de cardinalité, etc. OWL permet d'étendre RDF(S) en fournissant des primitives pour exprimer les notions qu'il n'était pas possible d'exprimer avec RDF(S). Ce langage dispose d'une sémantique formelle et fournit un support pour les inférences grâce à son lien avec les logiques de description.

ChapitreIV

le travail

sur

application

IV.1. Introduction

Après avoir établi une étude complète sur les ontologies ce chapitre est consacré à l'implémentation de l'application de ce dernier tout en présentant les langages et les outils utilisés ainsi que les copies d'écran de chaque étape.

IV.2. Outils et langages utilisés

Il existe plusieurs outils et langages pour la gestion des ontologies médicales. Parmi ces outils nous avons utilisé NetBeans IDE et Jena. Et parmi les langages nous avons utilisés JAVA.

IV.2.1. NetBeans

NetBeans est un environnement de développement intégré (EDI), placé en open source en juin 2000 sous licence CDDL (Common Dévelopment and Distribution License). En plus de Java, NetBeans permet également de supporter différents autres langages, comme C, C++, JavaScript, XML, et HTML. Il comprend toutes les caractéristiques d'un IDE moderne (éditeur en couleur, projets multi-langage, éditeur graphique d'interfaces et de pages Web).

Conçu en Java, NetBeans est disponible sous Windows, Linux, ou sous une version indépendante des systèmes d'exploitation (requérant une machine virtuelle Java).

NetBeans constitue par ailleurs une plate-forme qui permet le développement d'applications spécifiques (bibliothèque Swing (Java)). L'IDE NetBeans s'appuie sur cette plate-forme, il s'enrichit à l'aide de plugins.

IV.2.2. Jena

Jena est un API java open source développé par le laboratoire de Hewlett- Packard permettant la lecture et la manipulation des ontologies décrites en RDFS ou en OWL.

La page d'accueil de Jena, illustrée ci-dessous, est http://jena.sourceforge.net/

Rapport-Gratuit.com

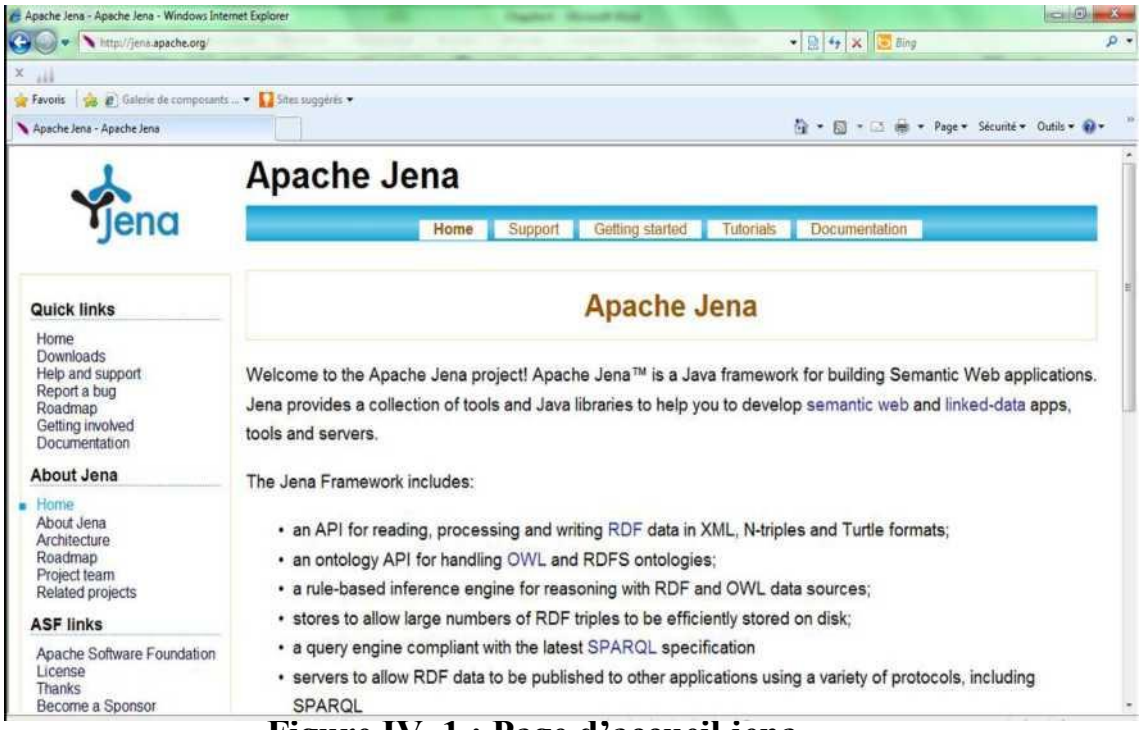

Figure IV. 1 : Page d'accueil jena.

IV.2.3. JAVA

Le langage Java est un langage de programmation informatique orienté objet créé par James Gosling et Patrick Naughton, employés de Sun Microsystems, avec le soutien de Bill Joy

Il existe bien d'autre IDE (Integrated DevelopmentEnvironement) pour le langage JAVA par exemple Eclips, JBuilder et NetBeans que nous avons utilisé…etc.

IV.3. Interfaces

Dans cette partie de ce chapitre, nous allons présenter les différentes formes de notre application.

L'écran ci-dessous présente l'interface principale de notre application qui a pour rôle de charger la fenêtre conforme à chaque utilisateur.

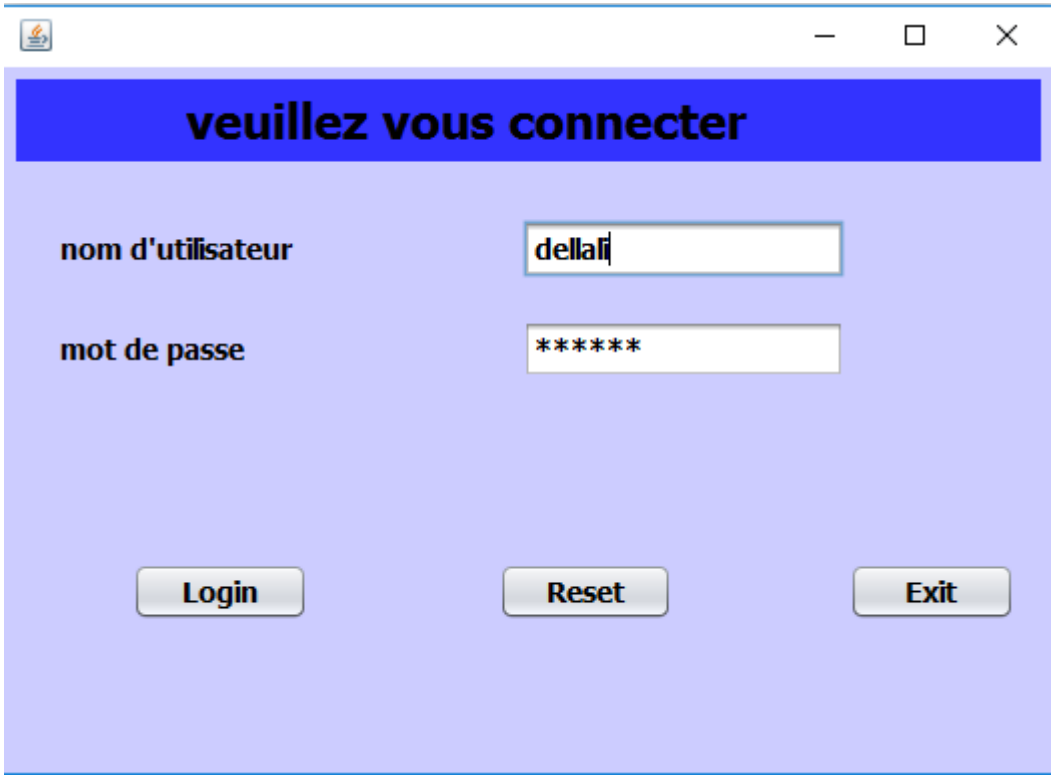

Figure IV.2 : Copie d'écran de l'interface principale.

La fenêtre sur la figure IV. 3 représente l'interface pour l'utilisateur médecin

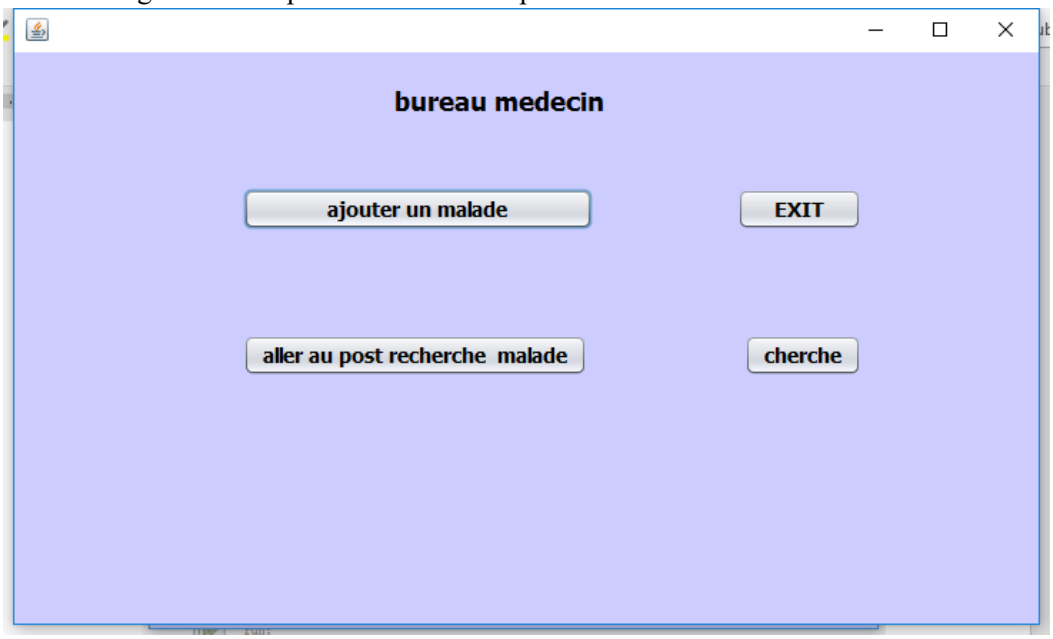

Figure IV. 3 : Copie d'écran de l'ouverture de bureau du médecin.

Cette fenêtre (Figure IV. 3) permet d'ajouter un malade à la liste de malade avec toutes ses coordonnées selon la figure IV.4.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

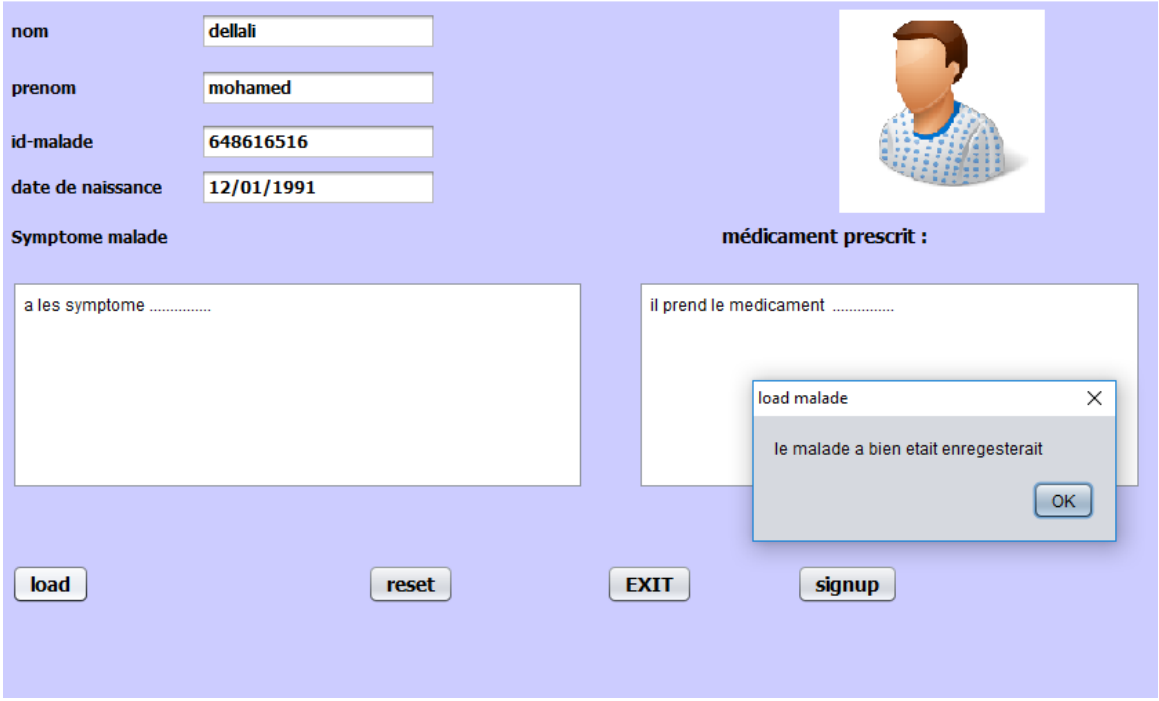

Figure IV.4 : Copie d'écran de l'ajout des malades.

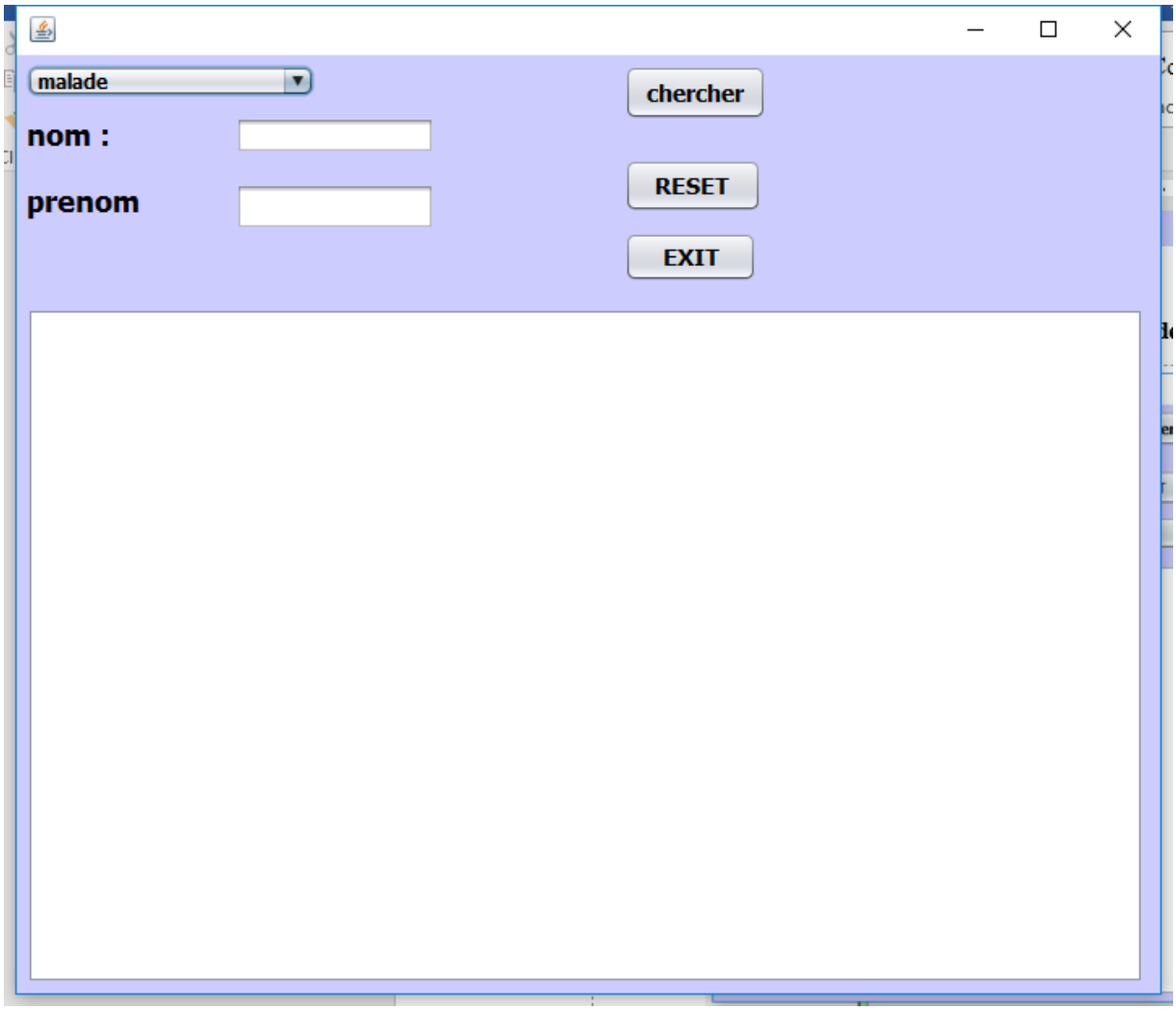

Figure IV. 5 :recherche détail malade

La figure IV.5 et par conséquent la figure IV.6, montrent le lien entre l'ontologie (figure IV.7)créée et le code Java qui permet de la manipuler.

```
User user = this.controller.login(this.username.getText(), this.password.getText());
if(user != null) {

      jTextPane1.setText(null);
  String queryString =
                "PREFIX ontologie:<http://www.co-ode.org/ontologies/ontologie.owl#> "
+               "PREFIX owl: <http://www.w3.org/2002/07/owl#> "
+               "PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> "
+               "PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>"
               +"PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>"


                + "SELECT  ?nom ?prenom ?date_naissace ?Symptome_malade ?medicament_prescrit"
        + " WHERE {?x  ontologie:id_malade  ?id_malade ; "
                +" ontologie:nom  'dellali'   ; "
                +"    ontologie:Symptome_malade ?Symptome_malade ;"
                +"    ontologie:medicament_prescrit ?medicament_prescrit ;"
                +"    ontologie:date_de_naissance    ?date_naissace ; "
                +"    ontologie:nom   ?nom ; "
          //          +"    ontologie:historique_du_malade   ?historique_du_malade ; "
                +" ontologie:prenom ?prenom   .}";
```

Figure IV. 6 : le code comp

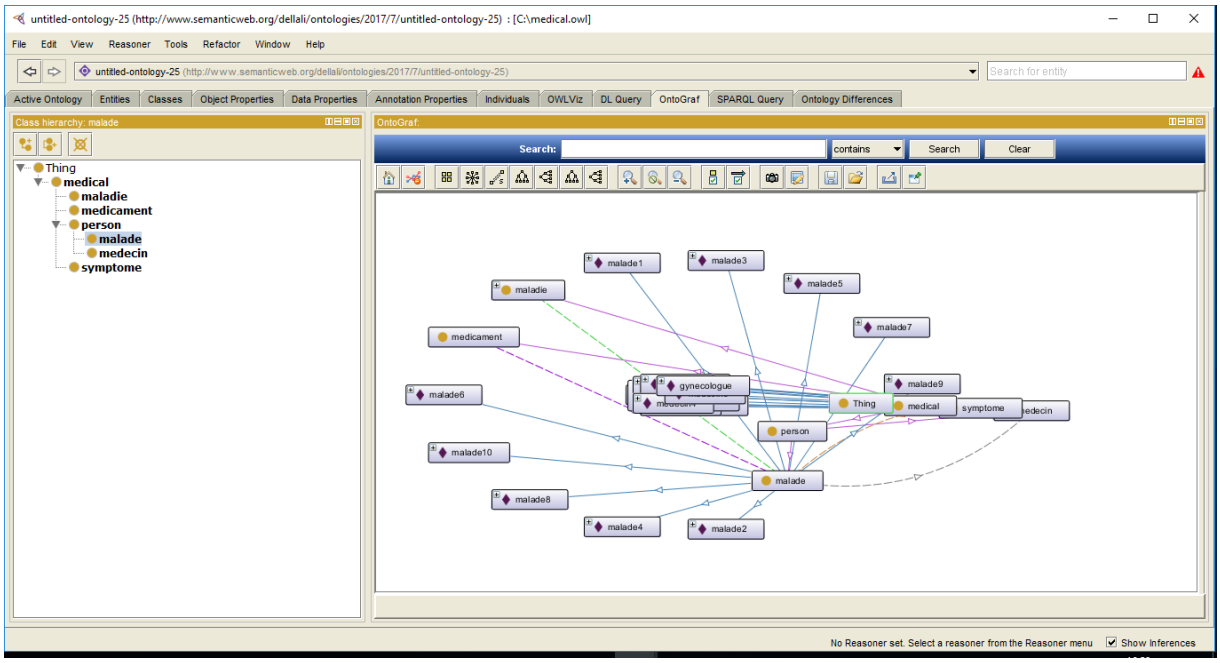

Figure IV. 7 : l'ontologie sur protégé

Ici est une copie d'écran de l'ontologie de Protégé

En cas que l'utilisateur est un patient l'application le dirige vers une autre fenêtre.

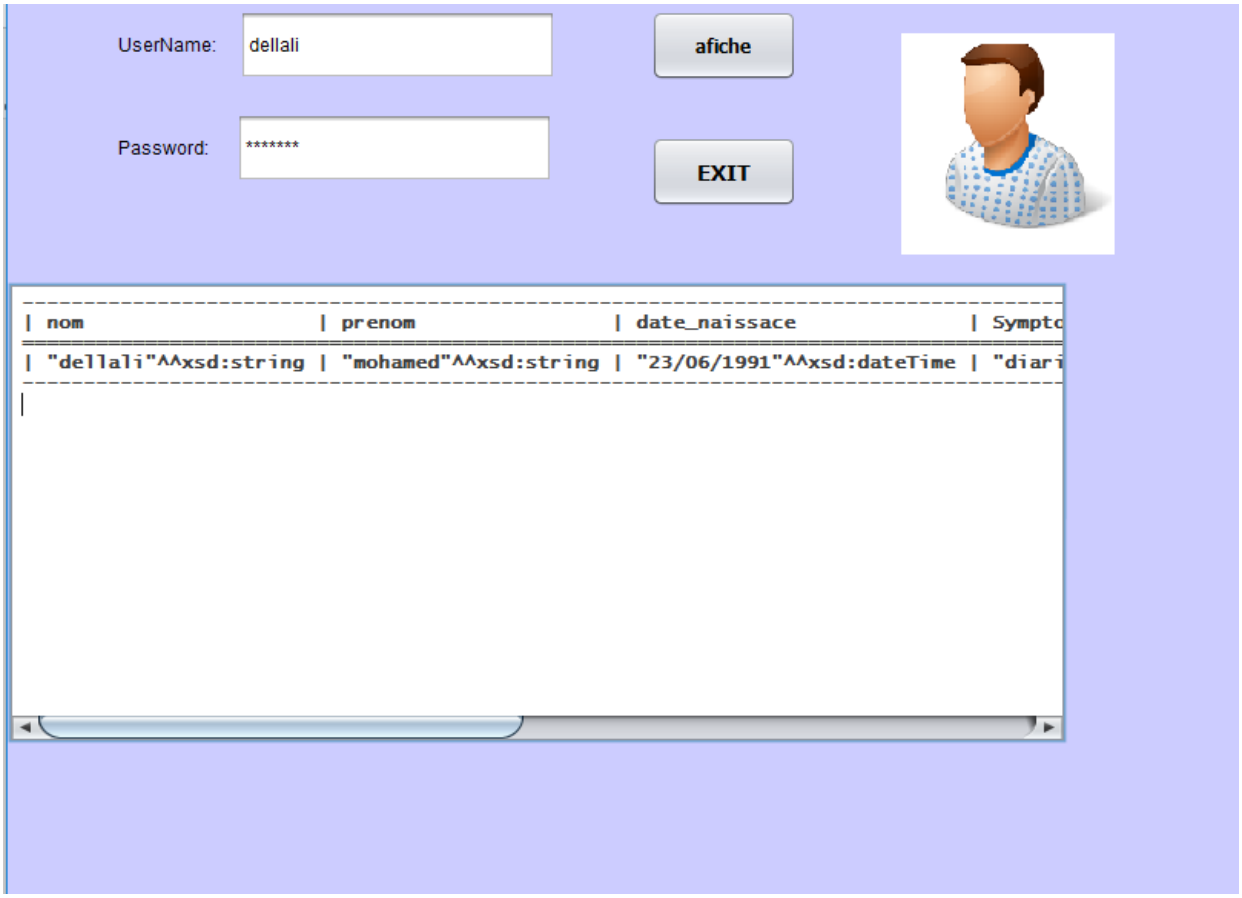

Figure IV. 8 : la fenêtre patient

En cas où l'utilisateur est un administrateur, l'application affiche une troisième fenêtre.

Figure IV. 9 : la fenêtre administrateur

Figure IV. 10 : la fenêtre ajout de coordonnées

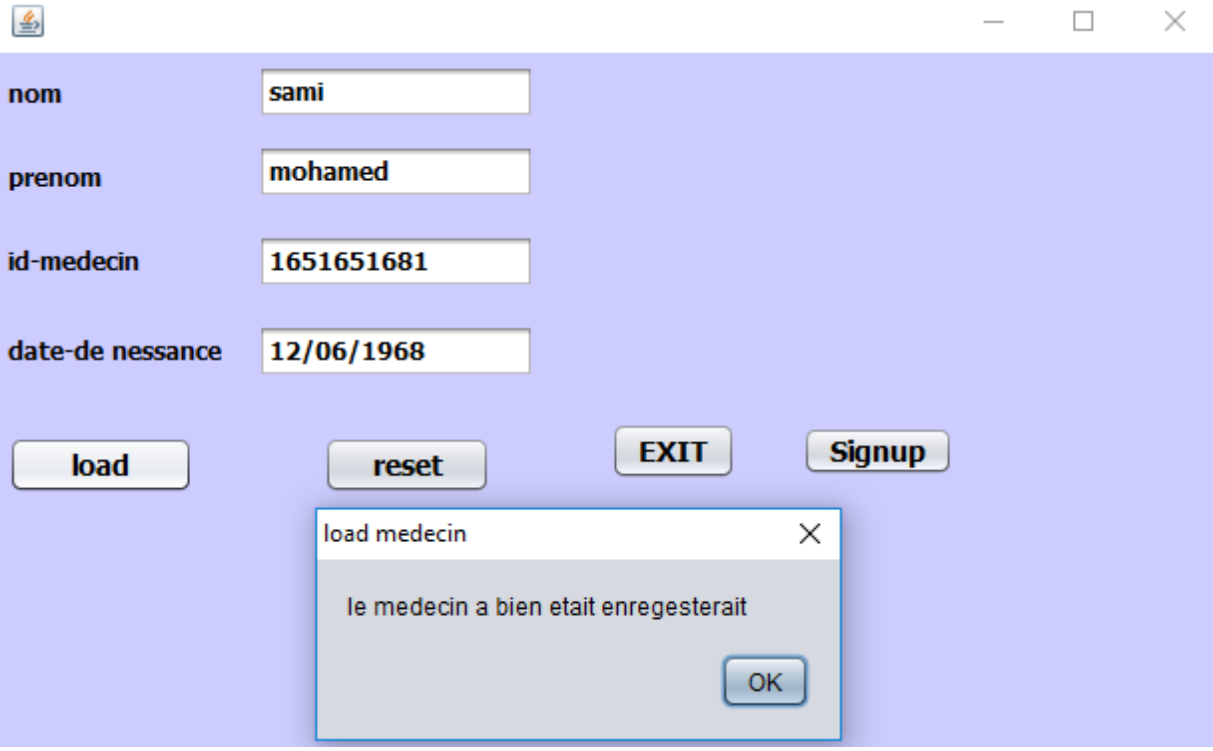

Figure IV. 11 : Ajout d'un médecin

IV.4. Conclusion

En utilisant les outils NetBeans IDE et Jena, et en utilisant le langage JAVA, nous avons développé une application de manipulation d'ontologie médicale. Les différentes étapes de la gestion de l'ontologie ont été développées dans ce chapitre. Après l'ouverture d'une ontologie, on définit les classes et les sous classes, ainsi que leurs créations ou suppressions. Nous avons aussi montré comment créer ou supprimer des concepts.

Conclusion Générale

Les ontologies médicales servent de plate-forme aux praticiens pour une prise de décision adéquate dans leur diagnostic. La prise en compte des différents aspects liés aux patients, aux maladies, aux remèdes et qui sont très nombreux est facilité par la création de telles ontologies médicales. Elle peut regrouper de nombreuses informations ainsi que les liens entres elles que ce soit du type hiérarchique (classes et sous classes), relations entre les concepts, propriétés des individus, etc...

Dans notre travail, nous avons développé un logiciel pour gérer les ontologies médicales, qui structure sous forme de graphes les différentes relations entre les patients avec toutes ses données, les maladies, les remèdes etc.., qui servira aux praticiens de rechercher l'information, de la traiter et de communiquer entre eux afin de donner le meilleur diagnostic.

Après avoir présenté les ontologies avec leurs caractéristiques, leurs constructions et leurs méthodes d'application, ainsi que les domaines d'application de ces ontologies. Nous avons fait un appariement entre les ontologies afin de sortir avec les meilleures correspondances.

Le processus d'extraction, représente l'étape de découverte de la sémantique correspondant à la modélisation de la sémantique, le processus de représentation et enfin le processus d'alignement qui permet d'aligner les ontologies concernées pour une analyse et traitement optimum des connaissances.

Le choix du langage d'implémentation joue un rôle important dans la construction de l'application donnée. Une liste d'outils de construction d'ontologie est présentée.

L'état de l'art que nous avons présenté concerne les notions liées aux ontologies ainsi que les langages et techniques qui permettent de représenter les ontologies utilisées dans le cadre du Web sémantique.

En utilisant les outils NetBeans IDE et Jena, et en utilisant le langage JAVA, nous avons développé une application de manipulation d'ontologie médicale. Les différentes étapes de la gestion de l'ontologie ont été développées. Après l'ouverture d'une ontologie, les classes et les sous classes, ainsi que leurs créations ou suppressions ont été définit. Nous avons aussi montré comment créer ou supprimer des concepts.

Rapport-Gratuit.com

Références Bibliographiques

[1] : Neeches, FininT ,Fikes R.E, Gruber T.R, Senator T et Swartou W.R. « Enabling technology for knowledge sharing » Al Magazine. Vol.12, no 3, 1993.

[2] : Gruber T. (1993). A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2), 199-220.

[3] : Guarino N. (1997b). Understanding, building and using ontologies. International J. Human-Computer Studies, 46, 293-310.

[4] : Borst W. N. (1997). Construction of Engineering Ontologies. Center for Telematica and Information Technology, University of Tweenty, Enschede, NL.

[5] : Studer et al. 1998, Studer R., Benjamins R. et Fensel D. (1998). Knowledge Engineering: Principles and Methods. Data Knowledge Engineering.

[6] : H. Sofia Pinto, J.P. Martins, Reusing Ontologies, aifbhermes.aifb.uni-karlsruhe.de/AAAI2000/ CameraReady/HPinto00.pdf.

[7] : Mr. BELABED, « Introduction aux ontologies », cours de l'intelligence artificielle master 1 système d'information et de connaissance.

[8] : BORG96, A. Borgida. "On the relative expressiveness of description logics and predicate logics". Artificial Intelligence, volume 82, number 1-2, pages 353-367, 1996.

[9] : Blas et al. 98, BLAZQUEZ M. FERNANDEZ M., GARCIA-PINAR J. M. & GOMEZ-PEREZ A., Building Ontologies at the Knowledge Level using the Ontology Design Environment, in Proceedings of the Banff Workshop on Knwolcdge Acquisition for Knwolcdgc-based Systems, 1998.

[10] : Dien et al. 01, DIENG R., CORBY O., GANDON F., GIBOIN A., GOLEBIOWSKA J., MATTA N. et RIB 1ERE M., «Méthodes et outils pour la gestion des connaissances : une approche pluridisciplinaire du knowledge Management ». Dunod, 2 edition. 2001.

[11] : Pere 99, Paterno F., Model Based Design and Evaluation of Interactive Application. Springer Verlag, 1999.

[12] : Lassila, O., et McGuinness, D. (2001). The Role of Frame-Based Representation on the Semantic Web. Technical Report KSL-01-02, Knowledge Systems Laboratory.Stanford University, Stanford, California.

[13] : BACHIMONT, B. (2000). Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances. In R. TEULIER, J. CHARLET, et P. TCHOUNIKINE (Ed.), Coordinateurs, Ingénierie des connaissances. Paris : Le Harmattan.

[14] : Gômez-Pérez, A., Femândez-Lôpez, M., et Corcho, O. (2004). Ontological Engineering (with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web).

Springer.

[15] : Mizoguchi, R., Vanwelkenhuysen, J., &Ikeda, M. (1995). Task Ontology for reuse of problem solving knowledge. In N. Mars (Ed.), Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing (KBKS'95) (pp. 46-57). IOS Press.

[16] : Van Heijst, G., Schreiber, A. T., &Wielinga, B. J. (1997). Using explicit ontologies in KBS development. *International Journal of Human-Computer Studies , 45***,** 183-292.

[17] : Guarino, N. (1998). Formal Ontology in Information Systems. In N. Guarino (Ed.), *1st International Conference on Formal Ontology in Information Systems (FOIS'98)* (pp. 3-15). IOS Press.

[18] : Gruber, T. R. (1993a). A translation approach to portable ontology specification. *Knowledge Acquisition, 5* (2), 199-220.

[19] : Psyché, V., Mendes, O., &Bourdeau, J. (2003). Apport de l'ingénierie ontologique aux environnements de formation à distance.

[20] : FÜRST, F. (2002). *L'ingénierie ontologique.* Rapport de recherche n 02-07, Institut de Recherche en Informatique, université de Nantes.

[21] : Uschold, M., &Grüninger, M. (1996). Ontologies: Principles, Methods and applications..*Knowledge Engineering Review , 11* (2), 93-155.

[22] : DE HOOG R. (1998). "Methodologies for building Knowledge Based Systems: Achievements and Prospects". Florida: Liebowitz J (ed) Handbook of Expert Systems, CRC Press Chapter 1, Boca Raton.

[23] : IEEE Standard Glossary of Software Engineering Terminology. IEEE Computer Society. New York Std 610.121990.

[24] : IEEE Standard for Developping Software Life Cycle Processes (1996). IEEE Std 1074-1995. IEEE Computer Society. New York.

[25] : FERNANDEZ M., A. GOMEZ-PEREZ et al. "METHONTOLOGY : from ontological art owards ontological engineering". In Proceedings of the Spring Symposium Series on Ontological Engineering (AAAI'97), AAAI Press, 1997.

[26] : BALAKRISHNAN H., FRANS KAASHOEK M., KARGER D., MORRIS R. and STOICA I. "Looking Up Data in P2P Systems", Communication of ACM, 2003, vol. 2, 46 p.

[27] : ARPIREZ J. C., CORCHO O., FERNANDEZ-LOPEZ M. and GOMEZPEREZ A. "WebODE : a scalable workbench for ontological engineering". In Y. Gill, M. Musen and J. Shavlik (Eds), Proceedings of the first International Conference on Knowledge Capture, K-CAP'01, Victoria, Canada, October 20-23, 2001 (New York: ACM Press), pp. 6-13.

[29] : Melle BOUARROUDJ Samia, « Raisonnement sur une ontologie enrichie par enrichie des règles SWRL pour la recherche sémantique d'images annotées », thèse de magister, (Ecole Doctorale en Informatique de l'Est -Pole ANNABA, 2009.

Rapport-gratuit.com
LE NUMERO 1 MONDIAL DU MÉMOIRES

[28] : GOMEZ-PEREZ A, FERNANDEZ LOPEZ M., CORCHO O. "Ontological Engineering". London : Springer, 2004, 403 p.

[30] : SURE Y., ERDMANN M., ANGELE J., et al. "OntoEdit: Collaborative ontology development for the semantic web", presented at International Semantic Web Conference 2002 (ISWC 2002), Sardinia, Italy, 2002.

[31] : NOY N. et McGUINNESS D. L. "Ontology Development 101: A Guide to Creating Your First Ontology". Technical Report KSL-01-05Stanford: Knowledge Systems Laboratory, mars 2001.

[32] : USCHOLD M. "Building Ontologies: Towards Unified Methodology", presented at the 16th Annual conference of the British Computer Society Specialist Group on Expert Systems, Cambridge, UK, 1996.

[33] : USCHOLD M., and GRUNINGER MICHAEL. "Ontologies: Principles, Methods and Applications". Knowledge Engineering Review, 1996, vol. 11, No 2.

[34] : GRUNINGER M., JINTAE LEE. "Ontology Applications and Design". In Communications of the ACM, 2002, vol. 45, No 2, pp. 39-41.

[35] : BERNARAS A, LARESGOITI I, CORERA J. "Building and reusing ontologies for electrical network applications. In Wahster W (ed) European Conference on Artificial Intelligence (ECAI'96). Budapest, Hungary. John Wiley and Sons, Chichester, United Kingdom, pp 298-302.

[36] : SWARTOUT R. P. B., KNIGHT K., RUSS T. "Toward Distributed Use of Large-Scale Ontologies". Ontological Engineering, 1997, pp. 138-148.

[37] : NOY N.F., FERGESON R.W., MUSEN M. A. "The knowledge model of Protege-2000: Combining interoperability and flexibility".

[38] : NOY N. et McGUINNESS D. L. "Ontology Development 101: A Guide to Creating Your First Ontology". Technical Report KSL-01-05Stanford: Knowledge Systems Laboratory, mars 2001.

[39] : GANGEMI A., PISANELLI DM., STEVE G. "An Overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies. Data & Knowledge Engineering, 1999, vol. 31, No 2, pp 183-220.

[40] : STOREY M. M. MARGARET-ANNE, SILVA JOHN, and BEST N. E. CASEY, FERGESEN RAY, NOY NATASHA, "Jambalaya: Interactive visualization to enhance ontology authoring and knowledge acquisition in Protégé", presented at Workshop on Interactive Tools for Knowledge Capture, K-CAP-2001, Victoria, BC, Canada, 2001.

[41] : webmaster-hub.com/publication/L-autre-semantique-Le-Web.html.

[42] : FARQUHAR A., FIKES R. & RICE J., Ontolinguaserver : a tool for collaborative ontology construction, in International journal of Human-Computer studies (46), pages 707-727, 2000.

[43] :ONTOLINGUA,OntolinguaHomePage,http://www.ksl.stanford.edu/software/ontolingua/, 2004.

[44] : OILED, OIL Editor Home Page, http : //oiled.man.ac.uk/, 2004.

[45] :PROTEGE2000,Protege2000OntologyEditorHomePage,http://protege.stanford.edu/, 2002.

[46] : NOY N & MUSEN M. A., Evaluating ontology-mapping tools : requirements and experience, in Proceedings of the Workshop on Evaluation of Ontology Tools (EON'2002) at EKAW'2002, 2002. [47]:ODE,OntologyDesignEnvironmentHomePage, http://www.swi.psy.uva.nl/wondertools/html/ODE.html, 2004.

[48] : ONTOEDIT, Ontology Editor Home Page, http ://www.ontoprise.de/com/, 2004.

[49] : Fellbaum, 1999, Fellbaum, C. (1999). WordNet : An electronic lexical database. Cambridge, Massashussets : MIT Press.

[50] : Roche, 2005, Roche, C. (2005). Terminologie et ontologie. LAROUSSE - Revue, 157:1_11. [Rosse et Mejino, 2003] Rosse, C. et Mejino, J. (2003). A reference ontology for bioinformatics : the foundational model of anatomy. J Biomed Inform., 36:478_500.

[51] : [Gamperet al. 1999], Gamper, J., Nejdl, W. et Wolpers, M. (1999). Combining ontologies and terminologies in information systems. Proc. 5th International Congress on Terminology and Knowledge Engineering, Innsbruck, Austria.

[52] : Nardiet Brachman, 2003, Nardi, D. et Brachman, R. J. (2003). An introduction to description logics.

[53] : Uscholdet King, 1995, Uschold, M. et King, T. (1995). Towards a methodology for building ontologies. In Proceedings IJCAI-95, Workshop on Basic Ontological Issues in Knowledge Sharing, Canada.

[54] : Lassilaet Swick, 1999, Lassila, O. et Swick, R. R. (1999). Resource description framework (rdf) model and syntax specification. World Wide Web Consortium W3C Recommendation 22 February 1999 / W3C /.

[55] : McBride, 2004, McBride, B. (2004). The resource description framework (rdf) and its vocabulary description language rdfs. Handbook on Ontologies, pages 51_66.

[56] : McGuinness et al. 2002, McGuinness, D. L., Fikes, R., Hendler, J. A. et Stein, L. A. (2002). Daml+oil : An ontology language for the semantic web. IEEE Intelligent Systems, 17(2) (5):72_80.

[57] : Lopez, 2001, Lopez, F. (2001). Overview of methodologies for building ontologies. IJCAI-99 Workshop on Ontologies and Problem-Solving Methods : Lessons Learned and Future Trends.Intelligent Systems, CEUR Publications 1999, 16(1).

[58] : Corcho et al. 2003, Corcho, O., Femândez-Lôpez, M., Gômez-Pérez, A. et Lopez- Cima, A. (2003). Building legal ontologies with methontology and webode. Pages 142_157.

[59] : ulMurshed et Ali, 2005, ulMurshed, A. et Ali, M. E. (2005). Evaluation and ranking of

ontology construction tools. Rapport technique Technical Report DIT-05-013, University of Trento.

[60] : Lenat et Guha, 1990, Lenat, D. B. et Guha, R. V. (1990). Building large knowledge based systems. Reading, Massachusetts :Addison Wesley.

[61] : Jarraret Meersman, 2002, Jarrar, M. et Meersman, R. (2002). Formal ontology engineering in the dogma approach. Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics (ODBase 02), LNCS 2519:1238_1254.

[62] : InterOp, 2006, InterOp, I. (2006). State of the art and state of the practice including initial possible research orientations, d8.1. Rapport technique, InterOp, Interoperability Research for Networked Enterprises Applications and Software.

[63] : Euzenat, 1995, Euzenat, J. (1995). Building consensual knowledge bases : context and architecture. Towards very large knowledge bases (actes2nd international conference on building and sharing very large-scale knowledge bases (KBKS), Enschede (NL), (10-13 avril) 1995) IOS press, Amsterdam (NL), pages 143_155.

[64] : Farquhar et al. 1995, Farquhar, A., Fikes, R., Pratt, W. et Rice, J. (1995). Collaborative ontology construction for information integration. Stanford University.

[65] : Benjamins et Fensel, 1998, Benjamins, R. et Fensel, D. (1998). The ontological engineering initiative-ka. Proceedings of the 1st International Conference on Formal Ontologies in Information Systems, FOIS'98, Trento, Italy, pages 287_301.

[66] : Simonetet al. 2006, Simonet, M., Patriarche, R., Bechlioulis, A., Bernhard, D., Diallo, G., Messai, R. et Gedzelman, S. (2006). Multilingual enrichment of an ontology of cardiovascular diseases. Computers in Cardiology (CINC'06) Proceedings.

[67] : Bodenreider, 2004, Bodenreider, O. (2004). The unified medical language system (umls) : integrating biomedical terminology. 32(Database-Issue):267_270.

[68] : Neal et al. 1998, Neal, P., Shapiro, L. G. et Rosse, C. (1998). The digital anatomist structural abstraction : a scheme for the spatial description of anatomical entities. J Am Med Inform Assoc. AMIA'98 Symp. Suppl., pages 423_427.

[69] : Pisanelliet al. 1998, Pisanelli, D. M., Gangemi, A. et Steve, G. (1998). An ontological analysis of the umlsmetathesaurus. AMIA 98 Annual Symposium.

[70] : Mathieu d'Aquin, Natalya F. Noy, *Where to publish and find ontologies? A survey of ontology libraries*, Web Semantics: Science, Services and Agents on the World Wide Web, Volume 11, March 2012, Pages 96–111

[71] : McCarthy, J. 1980. Circumscription: A form of non-monotonic reasoning. *Artificial Intelligence* 13(1-2):23-79.

[72] : mesurer et manager la qualité de service chap2 page 3

[73] : Kifer, Michael (2008). "Rule Interchange Format: The Framework

[74] : Thomas R. Gruber. A translation approach to portable ontology specifications. In R. Mizoguchi, editor, *Proceedings of the Second Japanese Knowledge Acquisition for Knowledge-Based Systems Workshop*, Kobe, 1992. To appear in *Knowledge Acquisition*, June

1993.
[75] : http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/ontosaurus_demo.html
[76] : Domingue, J., Motta, E. and Corcho Garcia, O. (1999) Knowledge Modelling in WebOnto and OCML: A User Guide. Available in PDF or gzipped PDF
[77] : 1. Gennari J, et al. The evolution of Protégé: An environment for knowledge-based systems development. Int. Journal of Human-Computer Interaction. 2003;58(1)
[78] : https://www.nlm.nih.gov/mesh/filelist.html
[79] : https://www.nlm.nih.gov/bsd/pmresources.html
[80] :*Benson, Tim (2012). Principles of Health Interoperability HL7 and SNOMED.*
[81] :http://searchhealthit.techtarget.com/definition/Unified-Medical-Language-System