

République Algérienne Démocratique et Populaire
Université Abou Bakr Belkaid– Tlemcen
Faculté des Sciences
Département d'Informatique

Mémoire de fin d'études

pour l'obtention du diplôme de Master en Informatique

Option: Système d'Information et de Connaissances (S.I.C)

Thème

Application de recherche d'emploi

Réalisé par :

- Baroudi Abderrazak

Présenté le 21 Juin 2015 devant le jury composé de :

- | | |
|----------------|-------------|
| - Smahi M. I | (Président) |
| - Benramdane D | (Encadreur) |
| - Halfaoui A | (Examineur) |
| - Khitri S | (Examineur) |

Remerciements

Avant tout, je remercie DIEU le tout puissant pour m'avoir aidé à réaliser ce travail.

En premier lieu, je tiens à remercier très chaleureusement mon encadreur Mme Berramdane, Professeure à la faculté des sciences, département d'informatique, université de Tlemcen, pour son aide fructueuse, de m'avoir aidé, orienté, conseillé et soutenu pendant toute la durée de ce travail.

Mes remerciements et mes reconnaissances à tous ceux qui ont participé de près ou de loin à la réalisation de ce mémoire.

Finalement, je remercie tous mes enseignants qui m'ont suivi le long de mes études.

Veillez agréer mes professeurs l'expression de mes sentiments très respectueusement dévoués.

Dédicaces

Avec l'aide de Dieu le tout puissant, j'ai pu achever ce travail que je dédie :

A mes très chers parents, leur amour, leur grande tendresse, leur compréhension et leur patience envers moi. Je ne saurais jamais comment exprimer mes sentiments pour avoir veillé sur mon éducation, pour leur soutien moral et matériel, jamais je ne peux les remercier assez de m'avoir donné le meilleur d'eux même.

A mes frères et sœurs.

Aux âmes de ma grande mère et mon grand-père.

A mes oncles et tantes.

A toute la famille BAROUDI.

A tous mes amis.

A mon encadreur M^{me} Berramdane et sa famille, je vous souhaite le bonheur.

A toute la promotion 2^{ème} année master SIC 2013/2014.

Table des matières

INTRODUCTION GENERALES3

1. CONTEXTE DE L'ETUDE ET PROBLEMATIQUE3

2. CONTRIBUTION4

3. PLAN DU MEMOIRE 4

Chapitre I : LE WEB SEMANTIQUES5

I.1 INTRODUCTION 6

I.2 AVANTAGES 7

I.3 DIFFERENCE ENTRE LE WEB TRADITIONNEL ET LE WEB Sémantique7

I.4 DEFINITION DU WEB SEMANTIQUE 8

I.5 OBJECTIF DU WEB SEMANTIQUE 8

I.6 PRINCIPE GENERALE 9

I.7 PRINCIPES TECHNIQUES 10

I.8 LES LANGUAGES DU WEB SEMANTIQUE 10

I.8.1 XML 11

I.8.2 RDF 12

I.8.3 OWL ET ONTOLOGIES 13

I.9 L'ARCHITECTURE DU WEB SEMANTIQUE14

I.10 LES CHAMPS D'APPLICATION 18

I.11 CONCLUSION20

Chapitre II : LES ONTOLOGIES21

II.1 INTRODUCTION 22

II.2 DEFINITION DE L'ONTOLOGIE 22

II.3 LES OBJECTIF DE L'ONTOLOGIE 23

II.4 LES COMPOSANTS D'ONTOLOGIE 23

II.5 LE CICLE DE VIE DES ONTOLOGIES 24

II.6 LES DIMENSIONS DE CLASSIFICATION 27

II.7 LES EDITEURS D'ONTOLOGIES 33

II.7.1 OIEd 34

II.7.2 OntoEdit 34

II.7.3 WebODE 35

II.7.4 DOE 35

II.8 CONCLUSION 36

CHAPITRE III: L'AUTOMATISATION DU RECRUTEMENT 37

III.1 INTRODUCTION 38

III.2 L'OBJECTIF DE NOTRE TRAVAIL38

III.3 LE SYSTEME PROPOSE 39

III.4 MATERIEL ET ENVIRONNEMENT 39

III.4.1 LE SYSTEME D'EXPLOITATION 39

III.4.2 LANGAGE DE PROGRAMATION39

III.4.3 LES LIBRIRIES40

III.4.4 RESSOURCE MATERIELLE 40

III.5 PRESENTATION UML 40

III.5 Principe de fonctionnement48

III.5.1 Les captures de l'application et l'explication de leurs fonctionnement48

CONCLUSION GENERALE ET PERSPECTIVES54

CONCLUSION55

BIBLIOGRAPHIE56

1. Contexte de l'étude et problématique

L'accès massif d'internet par les personnes, les institutions et les entreprises a changé radicalement la façon dont fonctionne le marché de l'emploi. De nos jours, des milliers de candidats mettent en ligne leur Curriculum Vitæ (CV), et les entreprises ou les institutions publient des profils de postes recherchés. Analyser automatiquement cette quantité d'informations est une tâche difficile à accomplir. Ceci est dû, d'un côté à la masse grandissante de CV reçus par les départements de ressources humaines, et d'un autre à l'énorme diversité de la présentation des CV. En particulier, dans certaines sections (identité, formation, expérience et compétences) et leur organisation.

Si on ne peut pas parler vraiment de documents « non-structurés », on peut les qualifier de « trop librement structurés », répondant à une structure conceptuelle propre à chaque individu et difficile à modéliser. Nous nous situons dans la double perspective d'emplois académiques et commerciaux. L'employeur est ici une institution (université, grande école, centre de recherche) ou une entreprise, et les candidats présentant des dossiers adaptés pour correspondre au mieux aux profils recherchés.

Une solution peut être trouvée dans le Web sémantique qui vise à enrichir les structures syntaxiques du Web actuel avec leur contenu sémantique. Par conséquent, l'exploitation de ces technologies dans le cadre du recrutement électronique semble être bénéfique, surtout pour assurer un rapprochement automatique entre les offres et les demandes d'emploi. Le principe consiste à expliciter le contenu de ces documents à l'aide d'éléments ontologiques formant un référentiel commun (ontologie), inspiré des parties communes entre ces documents. Ce référentiel peut être utilisé par un recruteur (ou un chercheur d'emploi) pour annoter son document avec des éléments pertinents. Les annotations résultantes peuvent être exploitées pour automatiser le processus de rapprochement entre son document et les CVs (offres d'emploi) disponibles.

2. Contribution :

Notre contribution dans ce contexte consiste à proposer une ontologie pour la modélisation du contenu sémantique des documents CVs/Offres d'emploi permettant leur rapprochement automatique ultérieur. Due à l'importance des compétences sous-jacentes à ces documents pour leur rapprochement, l'ontologie construite se concentre sur une modélisation explicite de la compétence. Le modèle adopté assure un certain compromis entre la simplicité et la pertinence, en distinguant l'identification des compétences acquises de celles requises et en explicitant la sémantique de la relation pouvant exister entre les différents thèmes des compétences. Ces documents sont annotés à l'aide d'éléments ontologiques formant un référentiel commun (ontologie), inspiré des parties communes entre ces documents (diplôme, expérience professionnelle, compétence, qualification personnelle, exigence).

3. Plan du mémoire :

Le mémoire est organisé comme suit :

Chapitre I : LE WEB SEMANTIQUE

Chapitre II : LES ONTOLOGIES

Chapitre III : L'AUTOMATISATION DU RECRUTEMENT : ce chapitre exprime le travail de ce projet.

CHAPITRE I

Le web sémantique

I.1 INTRODUCTION :

La prolifération des documents sur internet rend de plus en plus indispensable la possession d'un outil de recherche avancé qui permet d'obtenir des résultats pertinents.

Les moteurs de recherche traditionnels se basent sur l'occurrence d'un mot dans un document et sur la popularité des pages. Ainsi les résultats fournis par ces moteurs sont pollués par des milliers de documents qui ne concernent pas la notion recherchée.

Une approche plus pertinente serait donc de pouvoir rechercher un sens, une notion et pas seulement un mot. Il suffirait alors d'un mot pour récupérer les termes qui ont une relation logique avec ce mot, comme par exemple les synonymes et les antonymes. Les résultats seraient donc mieux ciblés et on obtiendrait certainement moins de résultats mais dont la plupart sont susceptibles d'être intéressants pour notre requête. A l'aide de la sémantique qui correspond à l'étude du sens des mots et des phrases et leurs relations afin d'en comprendre leur contexte et de détecter les fautes de sens qu'ils induisent.

Il faut noter que le problème de recherche de l'information est intimement lié à la structure des documents cibles de la recherche. La difficulté de la recherche sur la vision actuelle d'internet émane du fait que la plupart des documents sur le web sont en HTML. Or les documents HTML ont le défaut d'avoir un contenu non structuré.

Seule la mise en page est structurée en balises HTML prédéfinies. Mais ces balises ne portent aucune information sur le contenu du document. La structure du document et les informations sur les données contenues dans le document (métadonnées) détiennent un rôle primordial dans notre approche. Nous nous intéressons dans ce projet aux avantages que présente l'approche Web Sémantique par rapport à cette problématique.

Le Web sémantique est donc encore une vision pour le futur introduite par Tim Berners-Lee (inventeur du Web en 1989 et directeur du W3C) pour exprimer que le Web peut être fait plus intelligent et qu'on doit pouvoir appliquer encore plus de traitement automatique sur le Web mondial. Maintenant, le Web sémantique est vraiment un axe actif de recherche au W3C pour les nouvelles innovations du Web. Les utilisateurs du futur bénéficiant d'un Web sémantique pourront trouver plus exactement ce qu'ils veulent dans le Web toujours plus volumineux, ce qui leur est peu aisé dans le Web d'aujourd'hui malgré l'existence de beaucoup de moteurs de recherche performants. La sémantique correspond à l'étude du sens des mots et des phrases et leurs relations afin d'en comprendre leur contexte et de détecter les fautes de sens qu'ils induisent.

I.2 AVANTAGES :

Le Web sémantique, est la continuation logique du Web actuel. L'avantage premier de Web sémantique par rapport au Web sous sa forme actuelle est la présence de documents structurés selon un certain formalisme permettant le traitement automatique de ces documents par des machines. On arrivera donc à générer, traiter et échanger des documents grâce à des logiciels. La conception du document est donc une étape cruciale. L'utilisateur pourra en faire différents aperçus ou représentations à sa guise. De plus, les raisonnements et les relations logiques que nous mettons en place entre les différents documents pourront être automatisés par des moteurs d'inférence. Ainsi même des données non explicites mais déductibles de documents existants, seront générés automatiquement, sans intervention humaine.

Ceci est d'autant plus intéressant que la quantité de documents est grande. Pour parvenir à ces fonctionnalités, il faut d'abord structurer les documents. [28]

I.3 Différences entre LE WEB Traditionnel et LE WEB Sémantique

Web traditionnel	Web Sémantique
Web informel, contenus hétérogènes Le web ne respecte pas les critères de spécification du W3C	Web formalisé, contenus structurés Le web répond à des normes et spécifications
Web opaque L'information est noyée dans le flux de données	Web transparent L'information est visible, claire et accessible
Web quantitatif, contenus quelconques On table sur l'accumulation d'informations disponibles sur le web	Web qualitatif, contenus de valeur On exploite la valeur ajoutée des informations disponibles sur le web
Recherche de mots On ne s'intéresse qu'aux termes isolés	Recherche de concepts On s'intéresse au sens des mots, au contexte

Tab I.1 – différence entre le web traditionnels et le web sémantique

I.4 DEFINITION DU WEB SEMANTIQUE :

Le Web Sémantique est une vision du futur Web dans lequel l'information est donnée un sens explicite facilitant ainsi aux machines le traitement et l'intégration des informations sur le Web. Le Web Sémantique sera construit sur la capacité de XML de définir des schémas de balisage personnalisés et sur la flexibilité de l'approche RDF pour représenter les données. Si les machines sont supposées faire des tâches de raisonnement utiles sur ces documents, le langage doit aller au-delà des sémantiques de base du RDF Schéma. OWL a été conçu pour

répondre à ce besoin pour un langage d'ontologie pour le Web. OWL est une partie d'une pile évolutive des recommandations W3C relativement au Web Sémantique [32]

I.5 OBJECTIF DU WEB SEMANTIQUE :

Le Web sémantique a pour objectif de permettre à des machines d'accéder aux sources d'information et aux services du Web. Cette évolution pourra permettre d'automatiser la découverte des services pertinents par rapport à une requête, et également de combiner automatiquement plusieurs services lorsque la demande de l'utilisateur ne peut être satisfaite par un seul service prédéfini mais correspond à un service complexe. En effet, une description sémantique des ressources du Web permettra de les identifier de façon plus précise. Par ailleurs, il sera possible d'exécuter automatiquement les services du Web. Les travaux que nous effectuons se situent dans le cadre du Web Sémantique. Nous nous plaçons du point de vue de l'utilisateur final afin de l'aider à utiliser les services de commerce électronique accessibles via le Web, en lui proposant un système de médiation. Un système de médiation est une interface entre l'utilisateur et les services du Web d'un domaine d'application donné. Il doit donner l'impression à l'utilisateur d'utiliser un unique système alors que la satisfaction de sa demande peut exiger de composer plusieurs services délivrés par des fournisseurs différents. les deux objectifs principaux du web sémantiques sont :

- (i)Le Web doit devenir un moyen très puissant de coopération entre les humains.
- (ii)Les ordinateurs présents sur le réseau doivent coopérer : ils seront capables d'analyser toutes les données qui y circulent et de chercher les informations pertinentes pour l'utilisateur dans les bases de données accessibles aussi bien que dans les sites existants.

Ces données sont ensuite traitées et présentées correctement à l'utilisateur.

Il est important de noter que le terme "Sémantique" est employé dans un sens différent de celui utilisé dans le langage naturel. En effet, "Sémantique" signifie ici "interprétable par les machines". Les machines devront être capables d'utiliser des ressources provenant de diverses sources. C'est pourquoi ces ressources devront être disponibles, c'est-à-dire ouvertes et compréhensibles.

I.6 PRINCIPE GENERALE :

Sur le Web actuel, l'information est principalement présentée d'une façon qui la rend facile à envoyer et à montrer aux hommes. Les langages associés aux documents Web tels que

HTML et PDF sont des langages qui concernent la représentation visuelle des documents. Par exemple "<H2>Web Sémantique</H2>" n'apporte aucune information sur le contenu de la balise sans l'interprétation d'un lecteur humain. Toutefois, Internet n'est pas utilisé que par des humains, mais de plus en plus par des robots, des agents et d'autres algorithmes. Il est extrêmement difficile pour ces programmes d'interpréter l'information peu ou pas structurée disponible sur Internet.

L'objectif du Web Sémantique est d'être "un Web qui parle aux machines", c'est-à-dire dans lequel les ordinateurs offrent une meilleure aide aux utilisateurs de l'internet par le fait qu'ils se servent du contenu du Web. Aujourd'hui, Internet permet aux hommes d'accéder aux documents et services disponibles dans cet univers. Actuellement, les pages Web sont conçues par des humains pour que des humains les consultent. Le Web Sémantique a pour but d'augmenter le Web actuel avec de la connaissance formalisée et des descriptions sémantiques manipulables par des machines.

Cette ouverture de la connaissance aux ordinateurs permettra à ces derniers d'assister les hommes dans leurs tâches. Cette évolution du Web est considérée comme la troisième génération du Web. La contribution principale du Web d'origine a été de rendre des documents disponibles au travers de l'Internet, c'est-à-dire, en séparant la présentation d'un document de sa localisation. Ensuite, l'arrivée d'XML a permis de séparer la structure d'un document de sa présentation. Ce qui signifie qu'un document peut être perçu comme un conteneur de données dont la présentation peut varier selon l'utilisation. Finalement, l'idée du Web Sémantique consiste à séparer le sens de la structure, c'est-à-dire, étendre l'idée du document avec de la sémantique manipulable par des machines. Comme nous venons de le voir, l'interprétation et la manipulation sont facilitées par l'ajout de descriptions sémantiques, appelées métadonnées. Des métadonnées sont "des données sur des données". Le Web sémantique est entièrement fondé sur le Web "classique" et s'appuie sur sa fonction de base : un moyen de publier et consulter des documents. Seulement, les documents traités contiennent non pas des textes en langage naturel (français, anglais, etc.) mais des informations formalisées pour être traitées automatiquement.

Ces documents sont générés, traités et échangés par des logiciels. Il rend possible la compréhension de l'information aux ordinateurs. Le partage des connaissances se fait dans des langages essentiellement humains, le Web sémantique permet aux individus de s'exprimer dans des termes que les ordinateurs pourront également interpréter et échanger, et ce afin de résoudre des problèmes qui nous semblent fastidieux. [35]

I.7 PRINCIPES TECHNIQUES :

Le Web sémantique est fondé sur les protocoles et langages standards du Web : le protocole HTTP et les URI (Uniform Resource Identifier ou Identificateur Uniforme de Ressource). à ces standards s'ajoutent ceux qui sont propres au Web sémantique :

- (i)RDF : modèle conceptuel permettant de décrire toute donnée.
- (ii)RDF Schéma : langage permettant de créer des vocabulaires, ensembles de termes utilisés pour décrire des choses. [35]
- (iii)OWL : langage permettant de créer des vocabulaires complexes à des fins de traitements logiques poussés (inférences).

Ces trois standards sont ouverts et issus du W3C2. Ils forment l'ossature du Web Sémantique. [35]

I.8 LES LANGAGES DU WEB SEMANTIQUE :

Il a été bien reconnu dans la communauté du Web Sémantique que les ontologies jouent un rôle clé dans la livraison du Web Sémantique en facilitant le partage d'information entre les communautés d'humains et des agents logiciels. Les DTDs et les schémas XML peuvent être utilisés pour échanger les données entre les parties qui ont convenu les définitions avant l'échange. Cependant, leur manque de sémantique ne permet pas aux personnes de créer de nouveaux vocabulaires XML. Le même terme peut être utilisé avec différents sens dans différents contextes et des termes différents peuvent être utilisés pour des articles ayant le même sens. RDF et RDF Schéma ont essayé de résoudre ce problème en permettant d'associer des sémantiques simples aux identificateurs. Avec RDF Schéma, une personne peut définir des classes qui ayant plusieurs sous-classes et super classes, et peut aussi définir des propriétés pouvant avoir des sous propriétés, de domaines ou d'intervalle. Dans ce sens RDF Schéma est un langage d'ontologie pour le Web primitif. Pour atteindre l'interopérabilité entre les divers schémas développés de façon autonome, des sémantiques plus riches sont nécessaires.

Il semble clair que le Web Sémantique ne pourra voir le jour sans un minimum de standardisation. Différents consortiums et organismes mettent donc les acteurs autour d'une table pour définir les langages à utiliser dans le Web Sémantique. L'intérêt de cette approche est d'assurer des traitements uniformes sur des documents écrits dans ces langages. Les travaux de standardisation sont bien avancés. Durant le 10 février 2004, OWL et RDF sont devenus des recommandations du W3C. RDF est utilisé pour représenter l'information et pour échanger la connaissance sur le Web. OWL est utilisé pour publier et partager les ensembles de termes (appelées les ontologies), en supportant les recherches Web avancées, la gestion de la connaissance et des agents logiciels (W3C 2004).

I.8.1 XML :[44]

XML pour, extensible Markup Language, est un standard du World Wide Web Consortium. Tout comme le HTML, il est basé sur les balises, seulement en XML toutes les balises doivent être fermées. Cette rigueur facilite le traitement automatique d'un document. De plus, C'est un métalangage : il permet de décrire un vocabulaire et une grammaire associée selon un certain formalisme. L'intérêt premier de XML est de séparer le fond (le contenu) de la forme (le contenant ou la mise en page). Beaucoup de langages sont bâtis sur la syntaxe XML. On les appelle parfois dialectes XML. La structure d'un document XML est définie par une DTD (Document Type Définition).

La DTD peut être écrite dans un document à part puis référencée dans le document XML ou peut être directement intégrée dans ce dernier. On peut également écrire un fichier XML Schéma pour définir une structure plus détaillée et surtout pour compléter les définitions de types qui sont très restreintes en DTD. Nous définissons le type des éléments et les contraintes dans le XML Schéma. Une analyse peut alors vérifier si un document respecte bien les contraintes déterminées dans le schéma XML associé. Cependant, XML ne présente aucune information sur la sémantique des données exprimées dans le document. D'ou la nécessité d'utiliser RDF.

I.8.2 RDF :[42]

RDF, acronyme de Resource Description Framework, est un modèle de données. Par abus de langage il est commun de dire que c'est un " langage " d'assertion et d'annotation. C'est un des dialectes XML. RDF est un outil fondamental du Web Sémantique : il permet de définir des métadonnées afin de préciser les caractéristiques d'une information. Il établit des relations entre ressources. Il est donc particulièrement adapté aux annotations associées aux ressources du Web. Un document RDF est composé d'un ensemble de triplets < sujet, prédicat, objet >, chaque élément du triplet pouvant être un URI, un littéral ou une variable. Chaque triplet exprime un concept de base (statement en anglais). Grâce à ce regroupement par trois nous arrivons à un typage des objets que nous n'avons pas dans le web actuel. RDFS (RDF Schéma) n'est pas pour RDF ce que XML Schéma est pour XML. Il faut bien voir que XML Schéma contraint la structure d'un document XML, alors que RDF Schéma définit un vocabulaire utilisé dans les modèles RDF de données. Un document RDFS précise également les propriétés des différents objets modélisés, les domaines de valeurs possibles et décrit les relations entre ces différents objets. Un exemple de fichier RDF Schéma donnant les définitions des classes category et operator et la relation d'inclusion d'un opérateur dans une classe :

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rdf :RDFxmlns :rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns :rdfs="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#" >
```

```

<rdfs :Class rdf :ID="operator">
<rdfs :comment> The class of operators.
</rdfs :comment>
</rdfs :Class>
<rdfs :Class rdf :ID="category">
<rdfs :comment> The class of categories of operators .each operator belongs to a
category.
</rdfs :comment>
</rdfs :Class>
<rdfs :Property rdf :ID="belongsto">
<rdfs :domain rdf :resource= "#operator"/>
<rdfs :range rdf :resource= "#category"/>
</rdfs :Property>
</rdf :RDF> (exemple 1.1)

```

Une des restrictions fortes de RDF c'est l'utilisation des propriétés n'admettant que deux valeurs. Or cette restriction est trop forte car nous avons la plus part du temps plus de deux valeurs possibles. Par ailleurs RDF Schéma est fortement limité par la propriété de la hiérarchie

1.8.3 OWL ET ONTOLOGIES :[41]

Le OWL, acronyme de Web Ontology Language, permet d'étendre le vocabulaire et les propriétés de RDF. Il offre une grande souplesse dans la définition des relations.

Nous pouvons préciser qu'une propriété est l'inverse de l'autre, ce qui permet d'inférer des relations non explicites. Par exemple "A est le père de B " nous dit également que B est fils de A. Ceci paraît trivial, mais en informatique, il faut établir ce genre de raisonnement de base afin d'avoir des informations traitées "intelligemment ". Nous pouvons également établir des relations de disjonctions ou faire des unions. Voici le même (exemple 1.2) complété avec OWL :

```

<?xml version="1.0" encoding ="ISO-8859-1"?>
<rdf :RDF xmlns :owl = "http ://www.w3.org/2002/07/owl#" xmlns :rdf
= "http ://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns :rdfs =
"http ://www.w3.org/TR/1999/PR-rdf-schema-19990303#" >
<owl :Classrdf :ID="operator">
<owl :disjointWith rdf :resource ="#category"/>
<rdfs :comment> The class of operators.
</rdfs :comment>
</owl :Class>

```

```

<owl :Class rdf :ID="category">
  <rdfs :comment> The class of categories of operators. Each operator belongs to a
category.
</rdfs :comment>
</owl :Class>
<owl :ObjectProperty rdf :ID="belongsto">
  <rdfs :domain rdf :resource="#operator"/>
  <rdfs :range rdf :resource="#category"/>
  <owl :inverseOf rdf :resource ="#contains"/>
</owl :ObjectProperty>
<owl :ObjectProperty rdf :ID="contains ">
  <rdfs :domain rdf :resource="#category "/>
  <rdfs :range rdf :resource= "#operator"/>
  <owl :inverseOf rdf :resource ="#belongsto"/>
</owl :ObjectProperty>
</rdf :RDF>

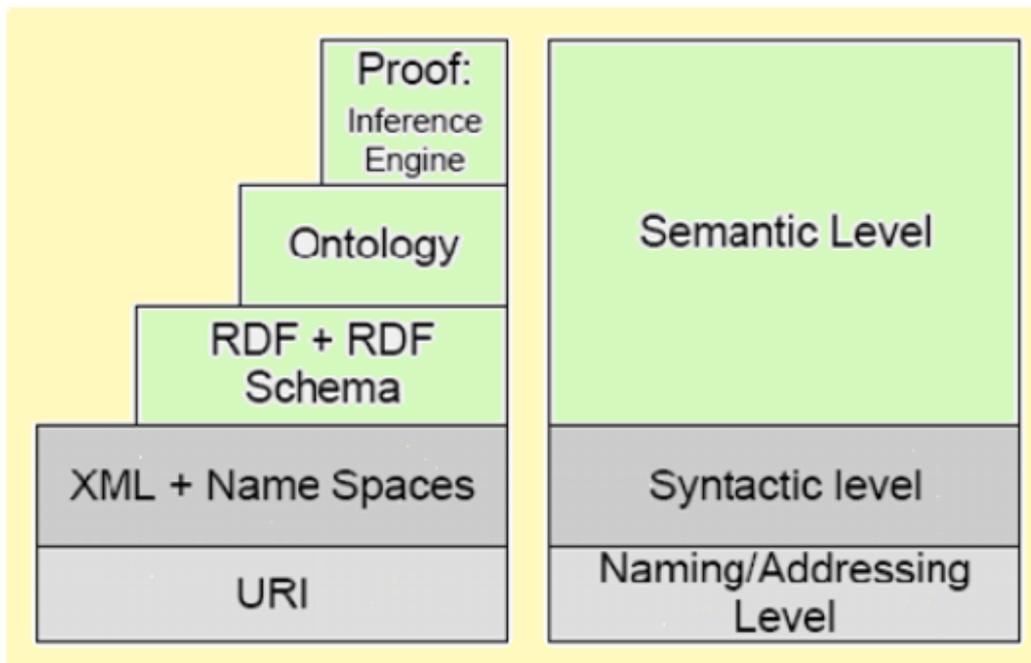
```

OWL met à disposition beaucoup d'outils, néanmoins des travaux sont en cours pour l'étendre d'avantage. Nous retrouvons dans la pile des couches du Web Sémantique suivante la démarche que nous avons suivie jusque là. Nous avons établi une structure XML et puis un modèle RDF dont la description est enrichie par des propriétés OWL.

Nous n'allons pas plus haut dans la pile pour le moment.

I.9 L'ARCHITECTURE DU WEB SEMANTIQUE:[37]

Le Web Sémantique n'est pas seulement une vision de l'avenir du Web. Certaines technologies sont déjà disponibles et figurent dans l'architecture illustrée dans la figure ci-dessous :



Figur

e I.1 – architecture du web sémantique

L'architecture repose sur des technologies de structuration de documents fondées sur XML. Sur cette base, des langages de métadonnées sémantiques de haut niveau ont été développés et permettent la description de ressources sur le Web. Afin de fournir un cadre interprétatif à ces métadonnées sémantiques, le Web Sémantique utilise des ontologies. Au niveau supérieur, le raisonnement sur les données est assuré par des mécanismes d'inférence, qui permettent d'une part de construire de la connaissance, mais aussi d'en maintenir la cohérence. Le Web Sémantique n'est pas l'Intelligence Artificielle présentée sous une autre forme. Le but n'est pas de développer des systèmes intelligents qui seraient aussi intelligents qu'un homme pour être utiles. Le point de référence n'est pas l'homme, mais le Web actuel qui est non sémantique. L'idée est d'ajouter progressivement des descriptions sémantiques de contenu sur le Web actuel afin de faire émerger petit à petit un Internet plus intelligent. Nous allons : maintenant présenter les trois niveaux importants du Web Sémantique Le niveau de l'adressage, le niveau syntaxique et le niveau sémantique.

– Le niveau Nommage/Adressage : Le World Wide Web repose sur un concept important qu'est l'URI (Uniform Resource Identifier). Tout ce qui est disponible sur Internet doit être identifié par un URI. Un URI identifie de manière unique et non ambiguë chaque

ressource du Web, comme une page, une adresse e-mail, ou une image. Le point central des URIs est l'URL (Uniform Resource Locator) traditionnelle utilisée pour définir les liens du Web (par exemple : `http://www-iasc.enst-bretagne.fr`). Ces URLs sont utilisées pour référencer des fichiers Web à travers un protocole particulier, comme HTTP ou FTP.

– Le niveau syntaxique : Le niveau syntaxique est le niveau de la structuration des documents. La spécification de la structure logique des documents repose sur XML. XML (eXtensible Markup Language) est un métalangage qui permet de définir d'autres langages. Les langages que définit XML sont des langages de présentation de documents. Il est aussi particulièrement adapté à l'échange de données. Son champ d'application est très vaste. XML permet de stocker des données structurées dans un fichier de texte. Elles y sont stockées dans un format simple et lisible relativement facilement. XML est utilisé sur le Web car il permet la description de documents électroniques par l'intermédiaire d'une DTD (Document Type Definition). Il fait le lien entre les documents et les logiciels qui les utilisent. En effet, le but de XML est de faciliter la diffusion et l'échange d'informations sur Internet. Les points forts d'XML en plus de sa simplicité sont :

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

L'avantage d'XML est la possibilité de personnaliser la présentation des documents en utilisant par exemple XSL (XML Stylesheet Language) qui permet de transformer automatiquement un fichier XML en une page HTML qui est consultable via un navigateur Internet. Son inconvénient : Il n'a pas de sémantique formelle permettant l'interprétation par la machine. XML décrit uniquement la structure de l'information, sa syntaxe.

– **Le niveau sémantique** : XML offre une solution à la structuration des données, mais la structure ne suffit pas à rendre les données complètement exploitables par un système.

Il est nécessaire de rendre les données interprétables par une machine par l'association d'une sémantique formelle. Un langage est interprétable par une machine s'il dispose d'informations sémantiques. Cela signifie que les symboles et la structure du langage font référence à un modèle sémantique. Le W3C propose d'ajouter des métadonnées et/ou des annotations permettant de décrire en terme de contenu chacune des ressources disponibles. La différence entre métadonnée et annotation réside dans le fait qu'une métadonnée serait attachée à une ressource et définie suivant un schéma, alors que l'annotation serait située au sein de la ressource et mise en place au cours d'un processus d'annotation.

– **RDF et RDFS** : Le meilleur langage à ce jour permettant d’apporter de la sémantique au Web est RDF (Resource Description Framework). Une description RDF est un graphe orienté étiqueté dont la syntaxe recommandée est fondée sur XML.

Il s’agit d’un formalisme utilisé pour représenter les propriétés d’une ressource et les valeurs de ces propriétés. RDF est donc un ensemble de triplets : Ressource, Propriété, Valeur. Une ressource peut être une page HTML, un élément situé dans une page HTML, un ensemble de pages, bref tout objet qui peut être identifié de façon unique par un URI. RDF peut être vu comme un langage déclaratif de représentation de connaissances.

La spécification de la syntaxe RDF est fondée sur XML, mais XML n’est pas le seul langage possible pour exprimer du RDF. En effet, le formalisme de RDF consiste en un ensemble de triplets. Mais de nombreux outils ont été développés autour d’XML. RDF ne propose pas d’outils permettant de spécifier le vocabulaire utilisé dans les descriptions. RDF Schéma (RDFS) est une extension de RDF permettant de décrire les concepts utilisés dans les descriptions et de définir des contraintes de type sur les objets et les valeurs des triplets. En fait, les triplets RDF sont les instances d’un RDF Schéma.

– **Les ontologies** : L’objectif principal du Web Sémantique concerne le partage et la réutilisation de l’information par tous : Hommes et Machines. Le Web Sémantique s’oriente vers des solutions qui consistent à ajouter de la connaissance à l’information de façon à améliorer la compréhension de cette dernière. L’ajout de la connaissance est obtenu soit par la gestion de métadonnées, soit par l’annotation des ressources. Un des problèmes soulevés porte sur la signification des concepts.

En effet, il est nécessaire qu’un même concept soit utilisé pour représenter la même chose dans l’indexation de différents documents. La solution est obtenue avec l’un des composants du Web Sémantique : les collections d’informations appelées ontologies. Le terme ”ontologie” est souvent utilisé aujourd’hui et possède de nombreuses définitions. Thomas Gruber [Gruber 1993] définit une ontologie de la façon suivante : ”Une ontologie est une spécification formelle explicite d’une conceptualisation partagée”. Les attributs ”formelle” et ”explicite” signifient qu’une ontologie permet une interprétation automatisée par la machine de la conceptualisation. En pratique, les ontologies sont des thésaurus et des hiérarchies de concepts évoluées et formalisées. Elles définissent des termes, des concepts, et leurs relations dans différents champs d’application. Elles diffèrent selon le domaine, l’utilisation ou les mécanismes de représentation utilisés. Voici quelques exemples d’ontologies :

1. Les ontologies pour les sciences (biologie, l’électronique, etc.)
2. Les ontologies pour le commerce (représentation des produits, etc.)
3. Les ontologies culturelles

4. Les ontologies de métadonnées

– **Logique, Preuve et Confiance** : Une ontologie définit les concepts d'un domaine particulier, mais n'explique pas comment il faut les utiliser. Par exemple, considérons un agent électronique qui doit être capable de déduire que si une commande est de 10 objets, que 4 ont été commandés à un premier fournisseur et que 6 ont été commandés à un second fournisseur, alors la commande est passée. Ce type de connaissance ne fait pas partie de l'ontologie mais concerne les actions, processus et déductions logiques sur les objets de l'ontologie. Le raisonnement sur la connaissance est donc un point important du Web Sémantique.

Un des inconvénients du raisonnement automatique est que l'utilisateur prendra pour acquis des conclusions qui pourraient provenir de faits erronés.

Il est nécessaire de savoir si l'on peut faire confiance aux informations disponibles.

La solution proposée est d'offrir des outils permettant d'évaluer et de commenter les informations disponibles sur le Web, puis grâce aux signatures électroniques, il sera possible d'identifier formellement la personne qui aura annoté le document, et par conséquent de s'assurer de la pertinence des commentaires.

I.10 LES CHAMPS D'APPLICATION :[34]

Les techniques du Web Sémantique, comme la description de ressources via des métadonnées et les ontologies peuvent être utilisées dans différents types d'applications telles que :

- Systèmes d'information
- Systèmes de gestion de connaissance
- E-Learning : l'objectif du E-Learning est de remplacer les anciennes façons concernant le temps, la place, le contenu de l'apprentissage prédéterminé avec des processus d'apprentissage à temps, à la place de travail, de manière personnalisée et à la demande de l'utilisateur. On trouve que les exigences principales du système du E-Learning sont la rapidité, le temps juste et l'apprentissage pertinent. Grâce à la propriété clé de l'architecture du Web sémantique avec le sens partagé commun, métadonnées traitables par les machines, ces exigences peuvent être satisfaites. Par exemple : le matériel d'apprentissage sera annoté de façon plus sémantique par une Ontologie et facilement combiné en un nouveau contenu du cours pour la demande de l'utilisateur. On peut le faire en utilisant les requêtes Web sémantique et la navigation à travers le matériel activée par un fond ontologique. Avec le Web sémantique, on peut avoir une plateforme adéquate pour implémenter un système E-Learning. On a besoin des moyens pour développer d'une ontologie d'apprentissage. Le

matériel d'apprentissage va être annoté par ontologie, et puis, on doit leur composer dans des cours et faire la livraison active des cours à travers des portails d'apprentissage.

– Commerce électronique c-à-d le e-commerce qui va nous permettre d'avoir un échange plus fluide d'information et de transactions entre tous les acteurs économiques, depuis l'offreur de produits ou les services jusqu'aux clients.

I.11 CONCLUSION

Aujourd'hui, personne ne peut nier le besoin d'y voir plus clair dans la recherche d'informations sur Internet. L'introduction du concept de portail communautaire tente de répondre à ce besoin. Les métadonnées, élément clé dans l'évolution des méthodes de recherche, dans l'évolution vers le Web sémantique, donnent lieu à d'intenses réflexions.

CHAPITRE II

Les Ontologies

II.1 INTRODUCTION :

L'humain est capable par la force de son langage de référencer deux contextes différents en utilisant le même terme, ou l'utilisation de plusieurs termes pour référencer la même idée. Or, l'automatisation n'est pas capable de faire ceci. Cela serait bien, si la machine pouvait comprendre les sens des termes, même s'ils émanent de plusieurs bases de données. Comment régler le cas où un terme prend deux significations différentes, comment lever toute ambiguïté "Ce problème est traité par le troisième pilier du SW et qu'on appelle ontologies" [45]

II.2 DEFINITION DE L'ONTOLOGIE :

Définition 1: Le terme "ontologie", selon le dictionnaire "petit Larousse 98" construit à partir des racines grecques ontos (ce qui existe, l'existant) et logos (le discours, l'étude), En informatique : l'ontologie est l'ensemble structure des termes et les concepts présentant le sens d'un champ d'information

Définition 2 : Plus simplement, l'ontologie est vue comme étant un vocabulaire qui a pour rôle de décrire et lier des modèles de connaissances [7].

Définition 3: Il s'agit de représentations formelles d'un domaine de connaissance sous la forme de terminologies dotées de relations sémantiques. [36].

Définition 4 : est un ensemble des objets reconnus comme existant dans le domaine.

Construire une ontologie c'est aussi décider de la manière d'être et d'exister des objets.

Dans cette définition, les objets ne sont pas pris dans un sens informatique mais comme objets du monde réel que le système modélise. Pour poursuivre vers une définition de l'ontologie, il nous semble indispensable de rappeler que les travaux sur les ontologies sont développés dans un contexte informatique - que ce soit celui de l'Ingénierie des connaissances, de l'Intelligence artificielle ou de la gestion et des systèmes d'information ou plus spécifiquement ici le contexte du Web Sémantique - où le but final est de spécifier un artefact informatique. Dans ce contexte, l'ontologie devient alors un modèle des objets existants qui y fait référence à travers des concepts, les objets du domaine.

Définition 5 : une ontologie est une spécification explicite d'une conceptualisation [11] "une conceptualisation est une vision abstraite et simplifiée du domaine à représenter.

Elle contient les objets, les concepts et les entités qui décrivent le domaine en plus les associations qui existent entre eux

II.3 LES OBJECTIF DE L'ONTOLOGIE : [46]

On utilise l'ontologie dans différents domaines : la représentation d'informations et de connaissances, l'intégration des systèmes d'informations, la spécification des systèmes, etc. Mais aussi dans :

– La communication : L'ontologie ne permet jamais que deux mots différents possédant la même sémantique.

L'interopérabilité : L'ontologie peut être considérée comme un pont ou une passerelle entre les différents systèmes. "Elle sert à définir le format d'échange entre les systèmes.

II.4 LES COMPOSANTS D'ONTOLOGIE :

Les connaissances traduites par une ontologie sont à véhiculer à l'aide des éléments suivants [10] : 1) Concepts ; 2) Relations ; 3) Fonctions ; 4) Axiomes ; 5) Instances, 6) les attributs

1. Les concepts : aussi appelés termes ou classes de l'ontologie, correspondent aux abstractions pertinentes d'un segment de la réalité (le domaine du problème), retenues en fonction des objectifs qu'on se donne et de l'application envisagée pour l'ontologie. Selon Gomez Pérez ces concepts peuvent être classifiés selon plusieurs dimensions :

- 1) niveau d'abstraction (concret ou abstrait) ;
- 2) atomicité (élémentaire ou composée) ;
- 3) niveau de réalité (réel ou fictif). [10]

2. **Les relations** : traduisent les associations (pertinentes) existant entre les concepts présents dans le segment analysé de la réalité. Ces relations incluent les associations suivantes :

- 1) Sous-classe-de (généralisation - spécialisation) ;
- 2) Partie-de (agrégation ou composition) ;
- 3) Associée-à ;
- 4) Instance-de, etc.

Ces relations nous permettent d'apercevoir la structuration et l'interrelation des concepts, les uns par rapport aux autres.

3. Les fonctions constituent des cas particuliers de relations, dans laquelle un élément de la relation, le nième (extrant) est défini en fonction des n-1 éléments précédents (intrants).

4. Les axiomes constituent des assertions, acceptées comme vraies, à propos des abstractions du domaine traduites par l'ontologie.

5. Les instances constituent la définition extensionnelle de l'ontologie ; ces objets véhiculent les connaissances (statiques, factuelles) à propos du domaine du problème.

6. Les attributs (slot/facette) : les slots sont des restrictions des concepts ou des classes. A chaque concept est associé un ensemble de slots et à chaque slots est associé un ensemble de facettes. [10]

II.5 LE CYCLE DE VIE DES ONTOLOGIES : [38]

Les ontologies sont des objets vivants, et chaque étape de leur cycle de vie pose des problèmes de recherche. Le cycle de vie inspire du génie logiciel est proposé dans [6] (**Figure II.1**) : détection des besoins, conception, gestion et planification, évolution, diffusion, utilisation, évaluation.

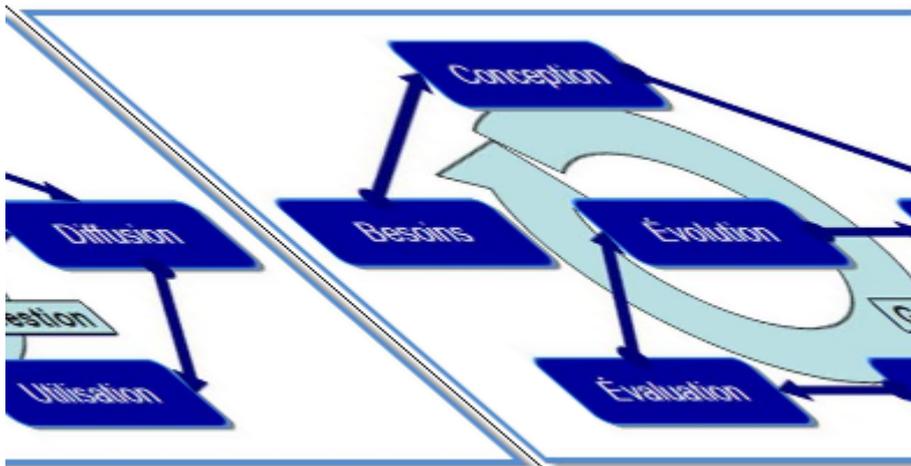


Figure II.1 – le cycle de vie de l'ontologie [10]

– **BESOINS ET EVALUATION** : L'activité de détection des besoins, préalable à la conception, et l'activité d'évaluation, lorsqu'une ontologie est utilisée, posent des problèmes

méthodologiques de recueil (analyse d'entretiens, questionnaires et sondages, étude de l'ergonomie et des usages) et d'identification (par exemple, modélisation par scénarios). En complément, la phase de détection des besoins demande un état des lieux initial approfondi, car elle ne peut reposer sur des études précédentes ou des retours d'utilisation, comme c'est le cas pour l'évaluation.

– **CONCEPTION ET EVOLUTION** : La phase de conception initiale et la phase d'évolution ont elles aussi en commun un certain nombre de problèmes:

1. spécification des solutions (conception participative, maquettage, prototypage) ;
2. acquisition des connaissances nécessaires (analyse de textes, traitement automatique de la langue naturelle, plateformes collaboratives) ;
3. conceptualisation et modélisation (design pattern ontologiques, méta ontologies, entretien avec les experts) ;
4. formalisation (méthodes et outils de l'Ontologie formelle, logiques de description et algorithmes de tableaux, analyse formelle de concepts, graphes conceptuels, formalismes du web sémantique RDF/S et OWL) ;
5. intégration de ressources existantes (alignement automatique d'ontologies, traduction) ;
6. implantation (graphes conceptuels, logiques de description, formalismes objets).

Notons aussi que l'évolution pose le problème de la maintenance de ce qui repose déjà sur l'ontologie. En effet, une ontologie est à la fois un objet vivant intéressant en soi et un ensemble de " primitives " pour décrire des faits du monde et des algorithmes sur ces faits. Lorsque l'ontologie change, ses changements ont un impact sur tout ce qui a été construit au-dessus. Le maintien de la cohérence dans une ontologie et au-dessus d'une ontologie, l'historique et la gestion des versions, la réingénierie et la propagation des changements après modification, sont des questions de recherche encore largement ouvertes. La maintenance de l'ontologie soulève donc des problèmes d'intégration technique et des problèmes d'intégration aux usages.

– **DIFFUSION** : La phase de diffusion s'intéresse au déploiement et à la mise en place de l'ontologie. Les problèmes de cette phase sont fortement contraints par l'architecture des solutions. Dans un contexte d'application web, on reposera sur des technologies idoines. Pour le partage de fichiers, des architectures paires à pair ou autres architectures distribuées peuvent être utilisées. Pour l'intégration d'applications, des architectures de services web peuvent être une solution. Dans toutes ces architectures (serveurs web, services web, pair à pair, agents, etc.) la distribution des ressources (données, modèles, applications et utilisateurs)

et leur hétérogénéité (syntaxes, sémantiques, protocoles, contextes, etc.) posent des problèmes de recherche sur l'interopérabilité (alignement et médiation) et le passage à l'échelle (largesbases, optimisation d'inférences, propagation de requêtes, syndication de données, composition de services, etc.).

– **LA PHASE D'UTILISATION** : regroupe toutes les activités reposant plus ou moins directement sur la disponibilité de l'ontologie, par exemple, l'annotation de ressources (traitement de la langue, rétro ingénierie de base de données, etc.), la résolution de requête (algorithme de projection de graphes avec contraintes), la déduction de connaissances et l'aide à la décision (moteurs d'inférence à base de règles), la navigation assistée et les services contextuels (analyse de contexte, identification et composition de services), l'analyse de gros volumes de connaissances (clustering, recherche de motifs récurrents). Toutes ces activités ont en commun de poser le problème de la conception des interactions avec l'utilisateur et de leur ergonomie (interfaces dynamiques). Sur ce point, l'ontologie apporte à la fois de nouvelles solutions (par exemple, les inférences exploitent les ontologies pour la génération dynamique d'éléments d'interfaces) et de nouveaux problèmes (par exemple, la complexification des modèles de données engendre des problèmes pour leur représentation et l'interaction avec ces représentations).

– **GESTION** : L'activité permanente de gestion et planification souligne qu'il est important d'avoir un travail de suivi et une politique globale pour détecter ou déclencher, préparer et évaluer les itérations du cycle et s'assurer que l'on reste dans le cercle vertueux des systèmes d'information

II.6 LES DIMENSIONS DE CLASSIFICATION :

Les ontologies peuvent être classifiées selon plusieurs dimensions. Parmi celles-ci, nous en examinerons quatre :

- 1) Objet de conceptualisation ;
- 2) Niveau de détail ;
- 3) Niveau de complétude ;
- 4) Niveau de formalisme de représentation.

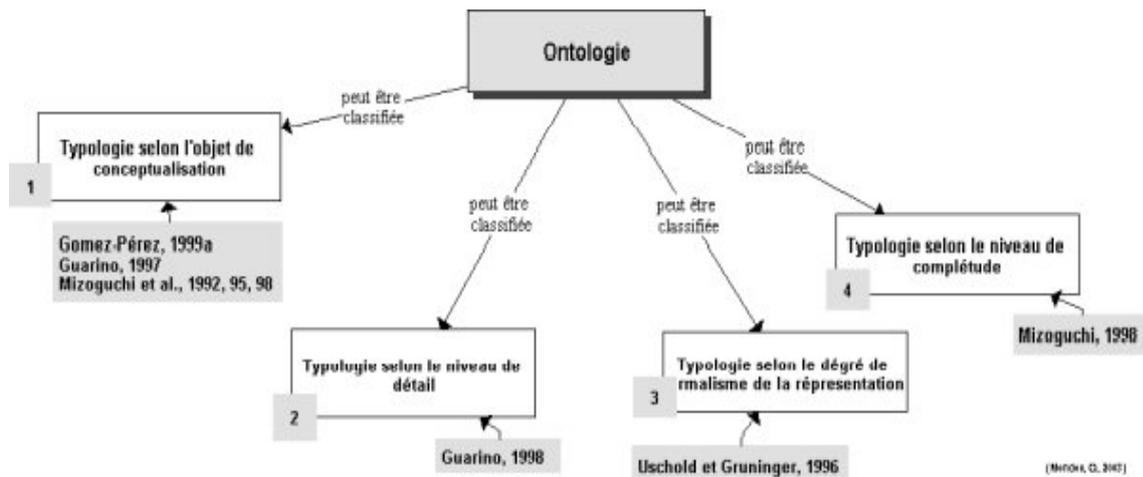


Figure II.2 – typologie de l’ontologie selon les dimensions de classification[10]

1. **TYPLOGIE SELON L’OBJET DE CONCEPTUALISATION** : Les ontologies classifiées selon leur objet de conceptualisation par [10],[17],[16],[24],[25], [26], sont de la façon suivante :

- 1) Représentation des connaissances ; 2) Supérieure/ Haut niveau ; 3) Générique ; 4) Domaine ; 5) Tâche ; 6) Application.

– Ontologie de représentation des connaissances/ [10], [24] ce type d’ontologies regroupe les concepts (primitives de représentation) impliqués dans la formalisation des connaissances. Un exemple est l’ontologie de Frame qui intègre les primitives de représentation des langages à base de frames : classes, instances, facettes, propriétés/slots, relations, restrictions, valeurs permises, etc.

– Ontologie supérieure ou de Haut niveau / [17], [20].

Cette ontologie est une ontologie générale. Son sujet est l’étude des catégories des choses qui existent dans le monde, soit les concepts de haute abstraction tels que : les entités, les événements, les états, les processus, les actions, le temps, l’espace, les relations, les propriétés. L’ontologie de haut de niveau est fondée sur : la théorie de l’identité, la méréologie (theory of whole and parts role) et la théorie de la dépendance. Guarino et Sowa ont poursuivi chacun indépendamment des recherches sur la théorie de l’ontologie. Tous deux intègrent les fondements philosophiques comme étant des principes à suivre pour concevoir l’ontologie de haut niveau ou supérieure. Sowa introduit deux concepts importants,

Continuant et Occurrent, et obtient douze catégories supérieures en combinant sept propriétés primitives. L'ontologie supérieure de Guarino consiste en deux mondes : une ontologie des Particuliers (choses qui existent dans le monde) et une ontologie des Universels comprenant les concepts nécessaires à décrire les Particuliers. La conformité aux principes de l'ontologie supérieure a son importance, lorsque le but est de standardiser la conception des ontologies.

– Ontologie Générique [10], [24].

Cette ontologie aussi appelée, méta-ontologies ou core ontologies, véhicule des connaissances génériques moins abstraites que celles véhiculées par l'ontologie de haut niveau, mais assez générales néanmoins pour être réutilisées à travers différents domaines. Elle peut adresser des connaissances factuelles (Generic domain ontology) ou encore des connaissances visant à résoudre des problèmes génériques (connaissances procédurales) appartenant à ou réutilisables à travers différents domaines (Generic task ontology). Deux exemples de ce type d'ontologies sont :

1) l'ontologie métréologique [5] contenant des relations, Partie-de et

2) l'ontologie topologique contenant des relations, Associé-à.

– Ontologie du Domaine [15] : Cette ontologie régit un ensemble de vocabulaires et de concepts qui décrit un domaine d'application ou monde cible. Elle permet de créer des modèles d'objets du monde cible. L'ontologie du domaine est une méta-description d'une représentation des connaissances, c'est-à-dire une sorte de méta-modèle de connaissance dont les concepts et propriétés sont de type déclaratif. La plupart des ontologies existantes sont des ontologies du domaine. Selon Mizoguchi, l'ontologie du domaine caractérise la connaissance du domaine ou la tâche est réalisée. Dans le contexte de la formation à distance, un domaine serait par exemple : le téléapprentissage.

– Ontologie de Tâches [15]. Ce type d'ontologies est utilisé pour conceptualiser des tâches spécifiques dans les systèmes, telles que les tâches de diagnostic, de planification, de conception, de configuration, de tutorat, soit tout ce qui concerne la résolution de problèmes. Elle régit un ensemble de vocabulaires et de concepts qui décrit une structure de résolution des problèmes inhérente aux tâches et indépendante du domaine. Selon [15], l'ontologie de tâche caractérise l'architecture computationnelle d'un système à base de connaissances qui réalise une tâche. Deux exemples d'utilisation de l'ontologie de tâche dans le domaine de l'éducation sont les suivants : 1) l'ontologie de formation par ordinateur - Computer Based Training Ontology [13] - qui régit un ensemble de concepts spécifiques à un système d'apprentissage inhérent à des ontologies de tâche ; et 2) l'ontologie des objectifs

d'apprentissage - Learning Goal Ontology [12] - qui décrit les Rôles des apprenants et des agents dans le cadre d'un apprentissage collaboratif.

– Ontologie d'Application. Cette ontologie est la plus spécifique. Les concepts dans l'ontologie d'application correspondent souvent aux rôles joués par les entités du domaine tout en exécutant une certaine activité [14].

2. PAR RAPPORT AU NIVEAU DE DETAIL : utilisé lors de la conceptualisation de l'ontologie en fonction de l'objectif opérationnel envisagé pour l'ontologie, deux catégories au moins peuvent être identifiées : 1) Granularité fine : correspondant à des ontologies très détaillées, possédant ainsi un vocabulaire plus riche capable d'assurer une description détaillée des concepts pertinents d'un domaine ou d'une tâche. Ce niveau de granularité peut s'avérer utile lorsqu'il s'agit d'établir un consensus entre les agents qui l'utiliseront ; 2) Granularité large : correspondant à un vocabulaire moins détaillé comme par exemple dans les scénarios d'utilisation spécifiques où les utilisateurs sont déjà préalablement d'accord à propos d'une conceptualisation sous-jacente.

Les ontologies de haut niveau possèdent une granularité large, compte tenu que les concepts qu'elles traduisent sont normalement raffinés subséquentement dans d'autres ontologies de domaine ou d'application. [9]

3. TYPOLOGIE SELON LE NIVEAU DE COMPLETUDE : Le niveau de complétude a été abordé par [16] et [2]. À titre d'exemple, nous décrivons la typologie de [2]. Ce dernier propose la classification sur trois niveaux suivante :

- **Niveau 1) Sémantique** : Tous les concepts (caractérisés par un terme/libellé) doivent respecter les quatre principes différentiels :

- 1) Communauté avec l'ancêtre ;
- 2) Différence (spécification) par rapport à l'ancêtre ;
- 3) Communauté avec les concepts frères (situés au même niveau) ;
- 4) Différence par rapport aux concepts frères (sinon il n'aurait pas lieu de le définir).

Ces principes correspondent à l'engagement sémantique qui assure que chaque concept aura un sens univoque et non contextuel associé. Deux concepts sémantiques sont identiques si l'interprétation du terme/libellé à travers les quatre principes différentiels aboutit à un sens équivalent.

– **Niveau 2) Référentiel** : Outre les caractéristiques énoncées au niveau précédent, les concepts référentiels (ou formels) se caractérisent par un terme/libellé dont la sémantique est

définie par une extension d'objets. L'engagement ontologique spécifie les objets du domaine qui peuvent être associés au concept, conformément à sa signification formelle. Deux concepts formels seront identiques s'ils possèdent la même extension (ex : les concepts d'étoile du matin et d'étoile du soir associés à Vénus).

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

4. TYPOLOGIE SELON LE NIVEAU DU FORMALISME : Par rapport au niveau du formalisme de représentation du langage utilisé pour rendre l'ontologie opérationnelle, [23] proposent une classification comprenant quatre catégories :

- 1) Informelles : ontologies opérationnelles dans un langage naturel (sémantique ouverte) ;
- 2) Semi-informelles : utilisation d'un langage naturel structuré et limité ;
- 3) Semi-formelles : langage artificiel défini formellement ;
- 4) Formelles : utilisation d'un langage artificiel contenant une sémantique formelle, ainsi que des théorèmes et des preuves des propriétés telles la robustesse et l'exhaustivité [10]. Selon Studer, "il y a différents types d'ontologie et chaque type remplit un rôle différent dans le processus de construction du modèle du domaine". [21]

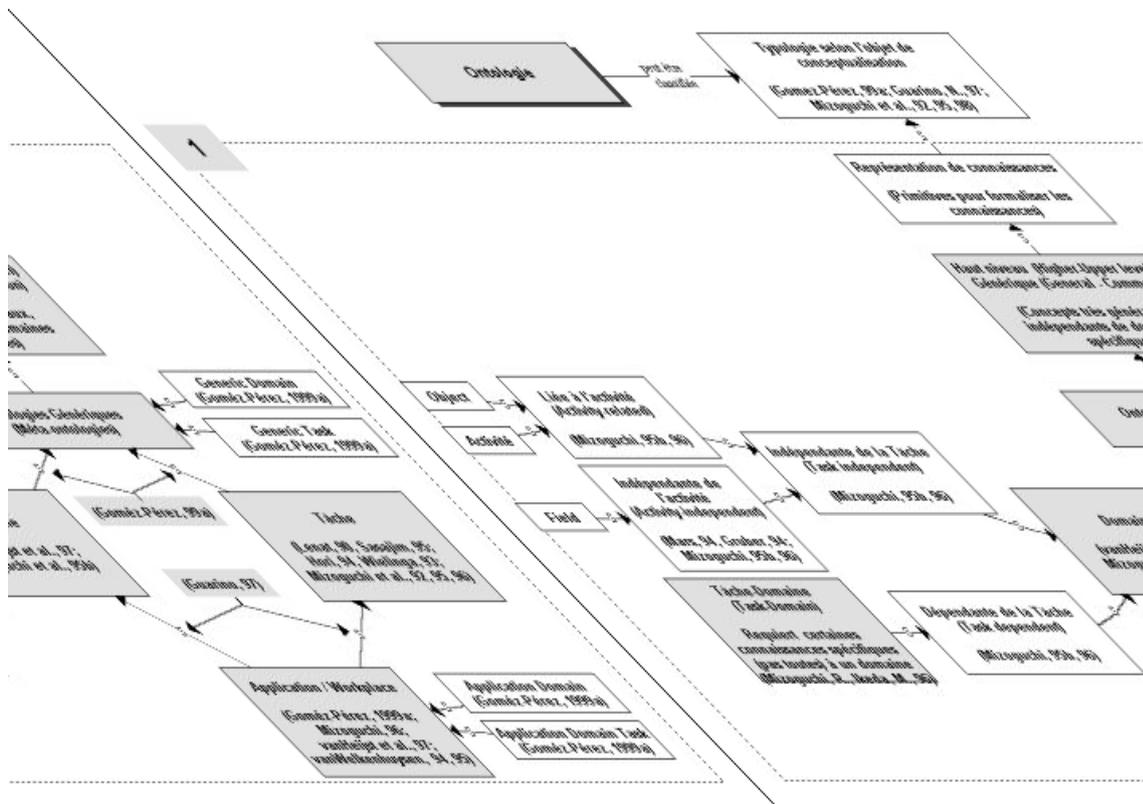


Figure II.3 –typologie selon l’objet de conceptualisation [21]

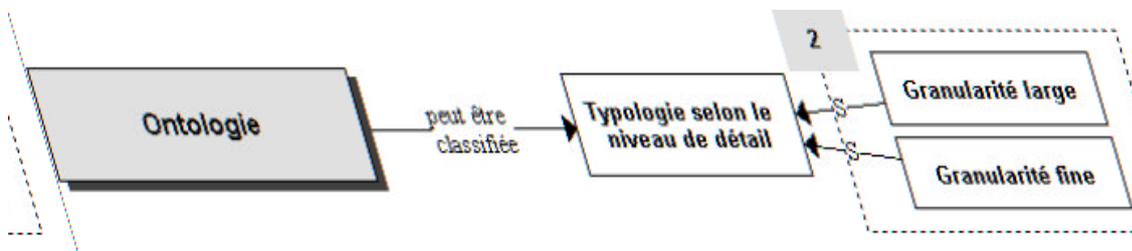


Figure II.4 – typologie selon le niveau de détail [10]

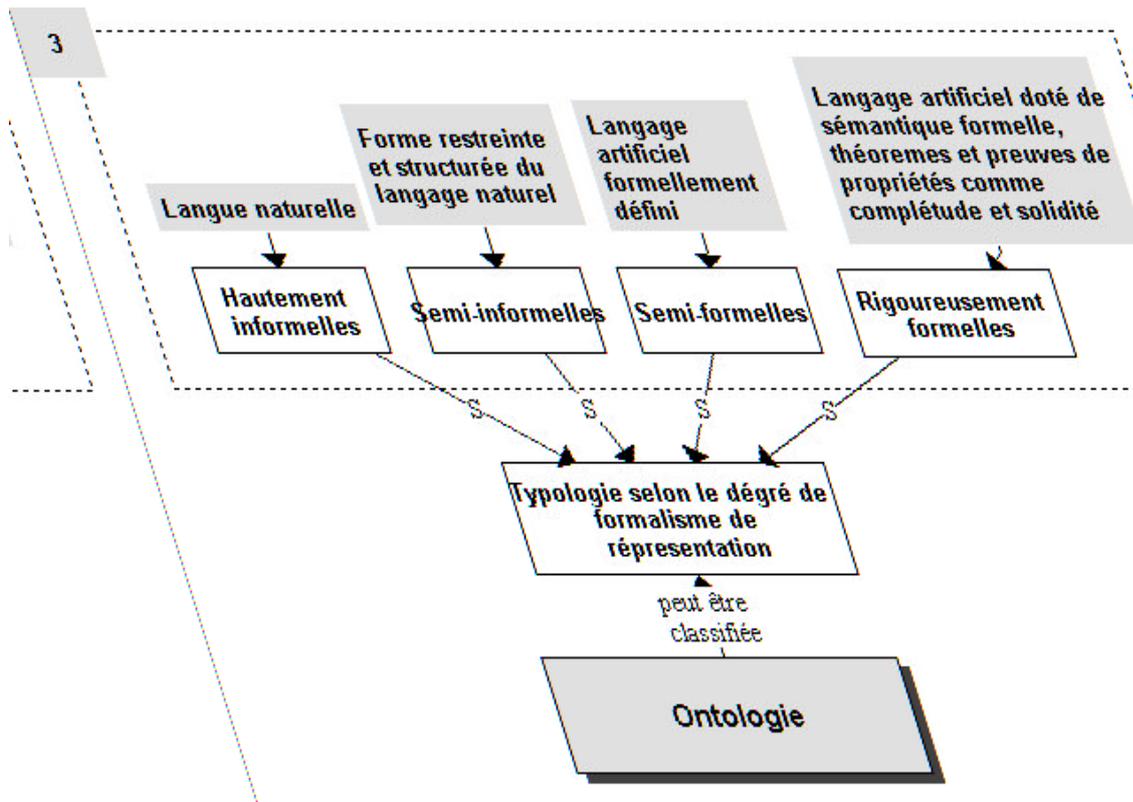


Figure II.5 – typologie selon le niveau de complétude [10]

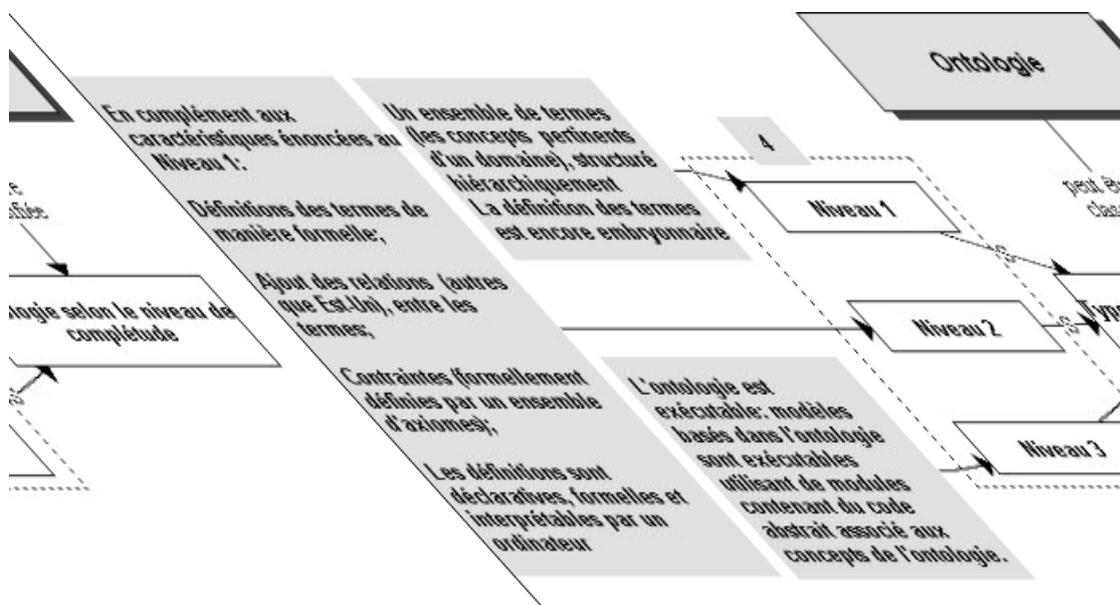


Figure II.6 – typologie selon le niveau du formalisme

– METHODES D'INGENIERIE ONTOLOGIQUE :

Les méthodologies recensées peuvent être analysées selon plusieurs critères ou dimensions, l'un d'entre eux étant le type du processus de construction:

1. Construction de nouvelles ontologies à partir de zéro ;
2. Réingénierie d'ontologies ;
3. Fusion ou intégration d'ontologies.

II.7 LES EDITEURS D'ONTOLOGIES :

De nombreux outils permettent aujourd'hui de construire des ontologies. Parmi ceux-ci, quelques-uns essaient de guider leur utilisateur dans l'élaboration de l'ontologie en suivant une méthodologie de conception plus ou moins complétée, en respectant des principes de cycle de vie et validation logiciels et les principes de l'IC. Dans tous les cas, force est de constater qu'aucun de ces outils n'a réussi à s'imposer et la réflexion sur l'outillage de la construction des ontologies reste donc ouverte. Les outils proposés peuvent se regrouper grossièrement en deux catégories. Dans la première, on trouve les plus anciens historiquement, qui permettent de spécifier les ontologies au niveau symbolique (voir, par exemple, le serveur ONTOLINGUA²⁶ [8]). Dans la seconde catégorie, les outils prennent mieux en compte l'importance du niveau des connaissances : ils proposent à leur utilisateur de créer l'ontologie de manière relativement indépendante de tout langage implémenté et prennent ensuite automatiquement en charge l'opérationnalisation de l'ontologie, en la transposant dans divers langages. Cette évolution tend à rapprocher les ontologies de leur but original : il semble en effet naturel de chercher à s'abstraire - dans un premier temps - du niveau symbolique si on veut obtenir une ontologie permettant un réel partage d'une compréhension. Cette dernière catégorie regroupe les outils principalement utilisés aujourd'hui.

II.7.1 OILEd :

OILEd développé sous la responsabilité de l'université de Manchester, a été conçu pour éditer des ontologies dans le langage de représentation OIL, un des précurseurs du langage OWL (Ontology Web Language) qui est aujourd'hui en voie d'être une recommandation W3C. Officiellement, il n'a pas d'autre ambition que de construire des

exemples montrant les vertus du langage pour lequel il a été créé. A ce titre, OILED est souvent considéré comme une simple interface d'un langage appartenant aux logiques de description. Néanmoins, il offre la plus grande partie de ce que l'on peut attendre d'un éditeur d'ontologies. On peut créer des hiérarchies de classes et spécialiser les rôles, et utiliser avec l'interface les types d'axiomes les plus courants. Cet éditeur offre également les services d'un raisonneur FaCT, qui permet de tester la satisfaisabilité des définitions de classes et de découvrir des subsomptions restées implicites dans l'ontologie.

II.7.2 OntoEdit :

Contrairement aux deux outils précédents, ONTOEDIT [22] n'est pas disponible gratuitement dans sa version complète. Il présente les fonctionnalités essentielles communes aux autres éditeurs (hiérarchie de concepts, expression d'axiomes, export de l'ontologie dans des langages divers) et a le mérite de s'appuyer sur une réflexion méthodologique significative. La modélisation des axiomes a fait l'attention de soins particuliers pour pouvoir être effectuée- en tout cas pour les types les plus répandus - indépendamment d'un formalisme privilégié et cela pour faciliter la traduction d'un langage de représentation à un autre. Il propose également une gestion originale des questionnaires de compétences. Un programme appelé ontokick permet la spécification des questions de compétence pour lesquelles l'ontologie doit fournir des réponses.

II.7.3 WebODE :

WebODE pour web Ontology Design Environment [1], développé par le LAI de Madrid, est une plateforme de conception d'ontologies fonctionnant en ligne. D'un point de vue méthodologique, l'outil fait suite à ODE, ce dernier 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.

Ces éditeurs assurent fidèlement le support de la méthodologie maison, en plus ils mettent l'accent sur la possibilité d'un travail collaboratif ou sur la mise à disposition d'outils complémentaires, comme un moteur d'inférences.

II.7.4 DOE :

Le dernier outil présenté ici est DOE pour Differential Ontology Editor [3] . Cet outil n'a pas pour ambition de concurrencer les grands environnements existants, mais plutôt de fournir un début d'implémentation à la méthodologie de structuration différentielle proposée

par B. Bachimont. A l'instar des autres éditeurs, il offre une représentation graphique des arbres de concepts et des relations de l'ontologie et permet d'interagir avec les hiérarchies. L'outil assiste également la saisie des principes différentiels issus de la méthodologie en automatisant partiellement cette tâche.

Le modèle de représentation de l'ontologie est finalement proche de celui du langage RDFS, à ceci près qu'il autorise la modélisation de relations n aires. Au niveau formel, l'éditeur est capable de faire quelques inférences en vérifiant la consistance de l'ontologie (propagation de l'arité le long de la hiérarchie des relations et héritage des domaines par exemple).

II.8 CONCLUSION:

Les ontologies apparaissent désormais comme une clé pour la manipulation au niveau sémantique et comme l'approche la plus prometteuse pour une possible automatisation.

Dans ce chapitre, nous avons commencé par introduire les ontologies nous avons passé ensuite à bien décrire les ontologies.

CHAPITRE III

L'AUTOMATISATION DU RECRUTEMENT

III.1 Introduction :

L'Internet s'est transformé rapidement en un outil de recrutement performant. Mais il connaît des faiblesses des outils dédiés à la gestion et à la recherche des CVs/offre d'emploi, l'écoulement de l'information sur ce marché de travail est loin d'être optimal.

Une solution peut être trouvée dans le Web sémantique qui vise à enrichir les structures syntaxiques du Web actuel avec leur contenu sémantique. Par conséquent, l'exploitation de ses technologies dans le cadre du recrutement électronique semble être

bénéfique, surtout pour assurer un rapprochement automatique entre les offres et les demandes d'emploi. Le principe consiste à expliciter le contenu de ces documents à l'aide d'éléments ontologiques formant un référentiel commun (ontologie), inspiré des parties communes entre ces documents (qualification personnelle, et exigence). Ce référentiel peut être utilisé par un recruteur (ou un chercheur d'emploi) pour annoter son document avec ses requis (ses acquis). Les méta-données résultantes peuvent être exploitées pour automatiser le processus de rapprochement entre son document et les CVs (offres d'emploi) disponibles. , nous proposons des techniques d'appariement sémantique entre CVs et offres d'emploi.

III.2 L'objectif de notre travail

Plusieurs travaux concernant le recrutement se basent sur la technique d'annotation sémantique, du fait de sa simplicité et de sa capacité à expliciter le contenu des documents. L'apport réside dans le fait que les données formalisées produites seront utilisées dans un rapprochement automatique entre ces documents.

Dans une première version de notre travail nous avons développé une ontologie détaillée pour la modélisation du contenu des CVs et des offres d'emploi. Notre objectif de tout ça c'est de trouver les CVs les plus proches par rapport les offres publiées (recruteurs) ou le contraire (chercheurs d'emploi).

III.3 LE SYSTEME PROPOSE

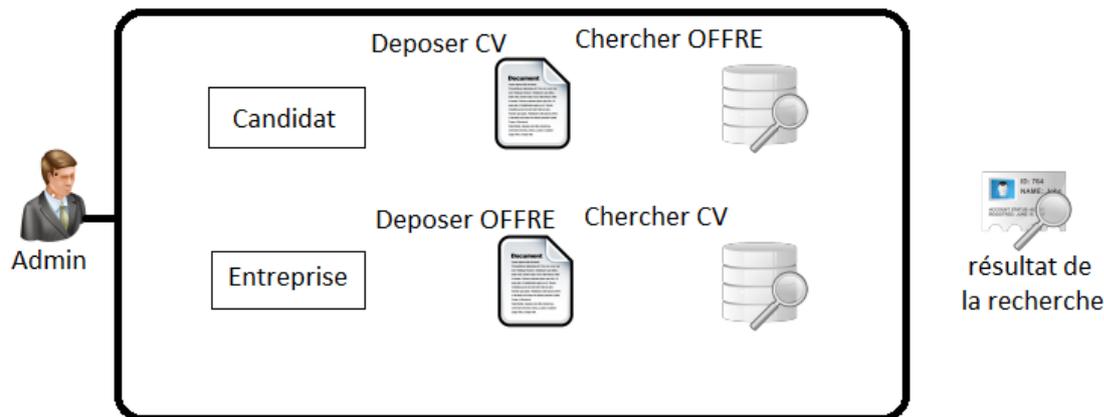


Figure III .1 architecture du système

III.4 Matériel et Environnement :

III.4.1 Le système d'exploitation :

L'environnement WINDOWS 8 a été choisi comme environnement de travail pour notre logiciel pour les raisons suivantes

- Une très bonne gestion de mémoire ;
- Une architecture orientée évènement ;
- Un graphisme indépendant des périphériques ;
- La notion de ressources.

III.4.2 Langage de programmation : On a travaillé sur la plate forme :

NETBEANS :[40]

Qui est un environnement de développement intégré (IDE) pour Java, placé en open source par Sun en juin 2000 sous licence CDDL (Common Development and Distribution License). En plus de Java, NetBeans permet également de supporter différents autres langages, comme Python, C, C++, XML et HTML. Il comprend toutes les caractéristiques d'un IDE moderne (éditeur en couleur, projets multi-langage, refactoring, éditeur graphique d'interfaces et de pages web). NetBeans est disponible sous Windows, Linux, Solaris (sur x86 et SPARC), Mac OS X et Open VMS. NetBeans est lui-même développé en Java, ce qui peut le rendre assez lent et gourmand en ressources mémoires.

– LANGUAGE DE MODELISATION UML : UML (Unified Modeling Language) est le nouveau concept de modélisation .Son développement a débuté en 1994.G.Booch et J.Rumbaugh deRational Software Corporation ont décidé de travailler ensemble pour réaliser une unification des principaux méthodes a objet

1. Les concepts : manipulés dans Uml sont de quatre natures. On distingue des concepts structurels(les classes, les interfaces, les collaboration, etc.), comportementaux(les interactions, les états d’objet), annotation(les notes),et de groupement(les packages, les sous-systèmes, etc.).

2. Les relations : permettent de lier les concepts. UML offre quatre types de relations: les associations, les généralisations, les dépendances et les réalisations.

3. Les diagrammes : constituent des représentations graphiques d’ensemble de concepts. Ils sont définis par des graphs connexes dont les sommets sont des concepts et les arcs des relations inter-concepts. Les diagrammes permettent de représenter les systèmes sous différentes perspectives et différent niveaux d’abstraction. UML offre neuf types de diagrammes offrant des visions statiques ou dynamiques des systèmes.

1. DIAGRAMME D’ACTIVITE : est une variante de diagramme d’états de transitions, organisé par rapport aux actions et principalement destiné à représenter le comportement interne d’une méthode (la réalisation d’une opération) ou d’un cas d’utilisation. Un diagramme d’activité représente l’état d’exécution d’une méthode ou d’un cas d’utilisation.

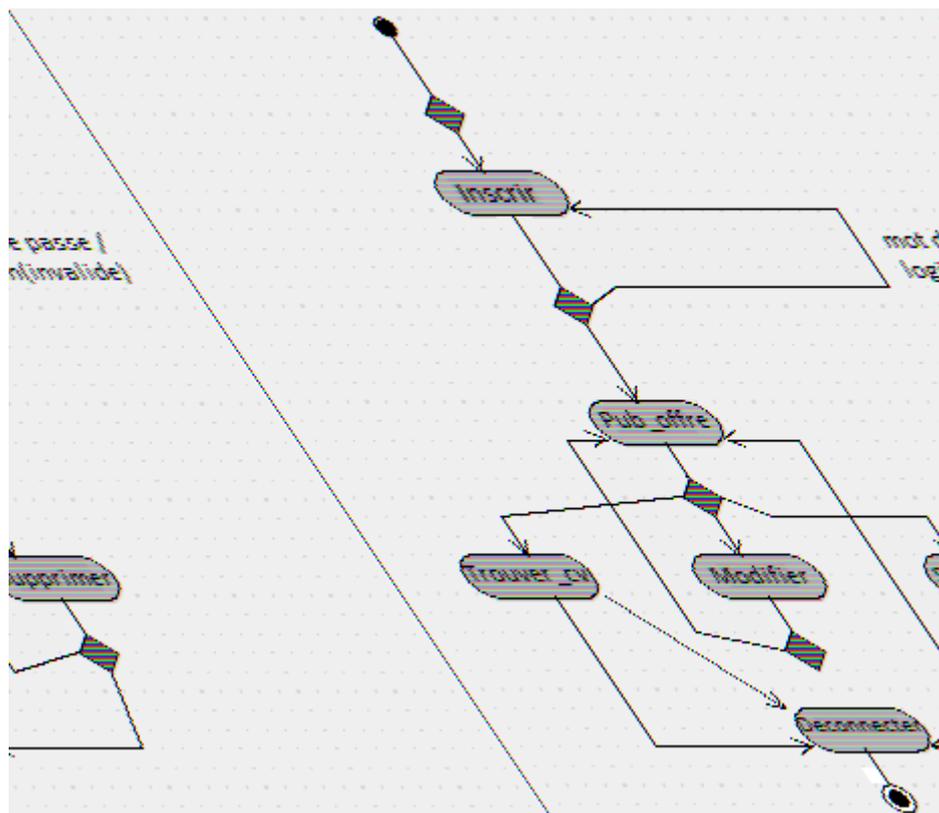


Figure III.3 – diagrammes d'activité d'administrateur

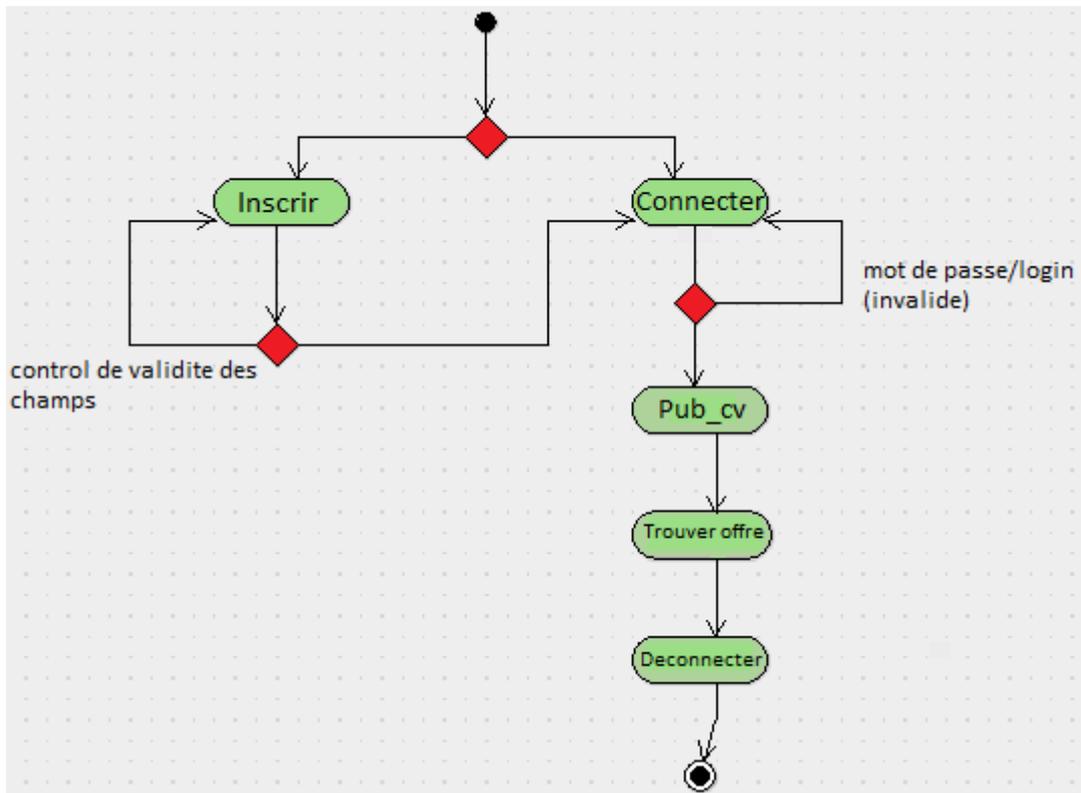


Figure III.4 – diagrammes d'activité de candidat

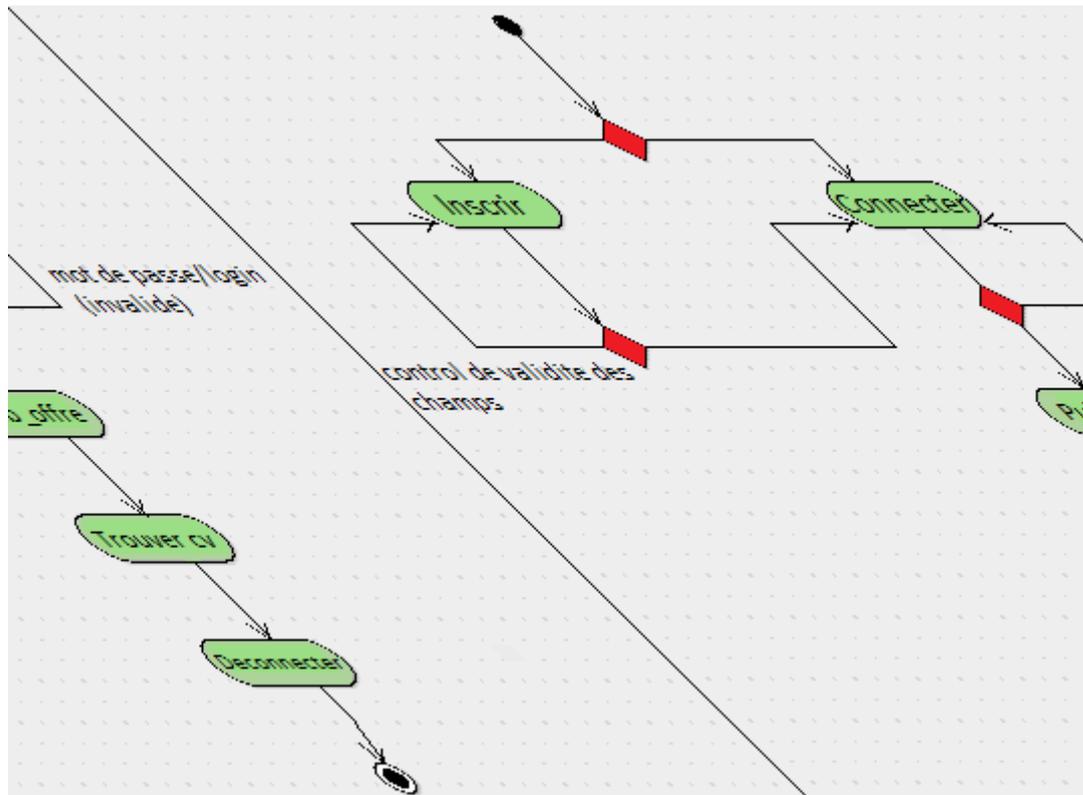


Figure III.5 – diagrammes d’activité de l’entreprise

2. DIAGRAMME DE CAS D’UTILISATION : permet de recueillir, d’analyser et d’organiser les besoins, et de recenser les grandes fonctionnalités d’un système. Il s’agit donc de la première étape UML d’analyse d’un système.

Un diagramme de cas d’utilisation capture le comportement d’un système, d’un sous-système, d’une classe ou d’un composant tel qu’un utilisateur extérieur le voit. Il scinde la fonctionnalité du système en unités cohérentes, les cas d’utilisation, ayant un sens pour les acteurs.

Les cas d’utilisation permettent d’exprimer le besoin des utilisateurs d’un système, ils sont donc une vision orientée utilisateur de ce besoin au contraire d’une vision informatique. Il ne faut pas négliger cette première étape pour produire un logiciel conforme aux attentes des utilisateurs.

Pour élaborer les cas d’utilisation, il faut se fonder sur des entretiens avec les utilisateurs.

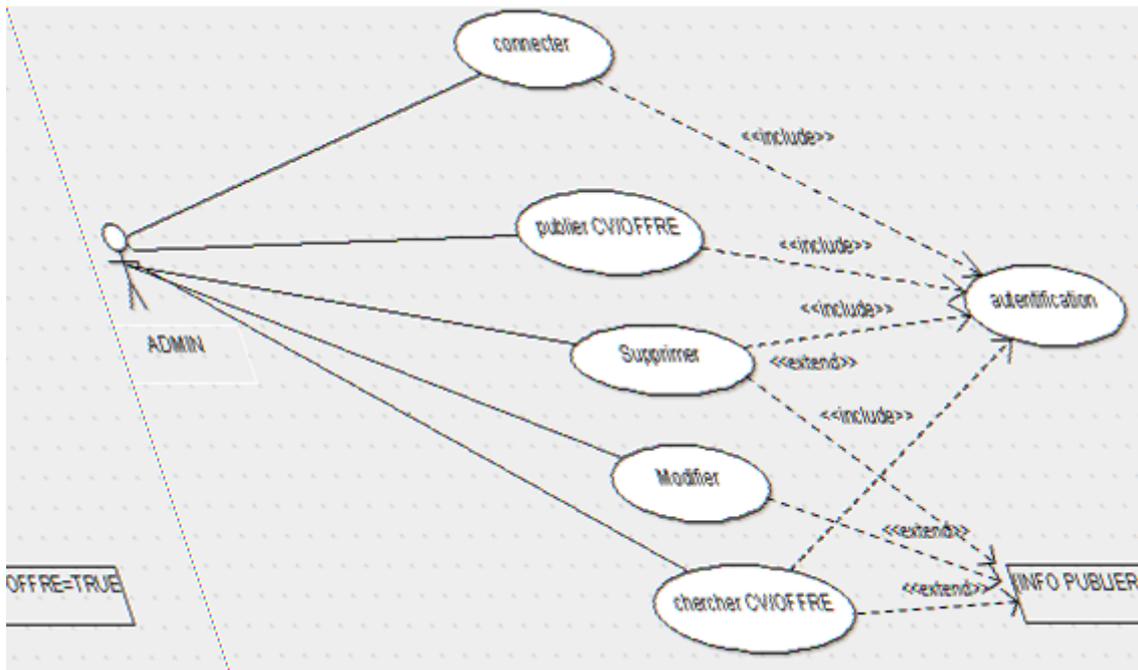


Figure III.6 – diagrammes de cas d'utilisation de l'administrateur

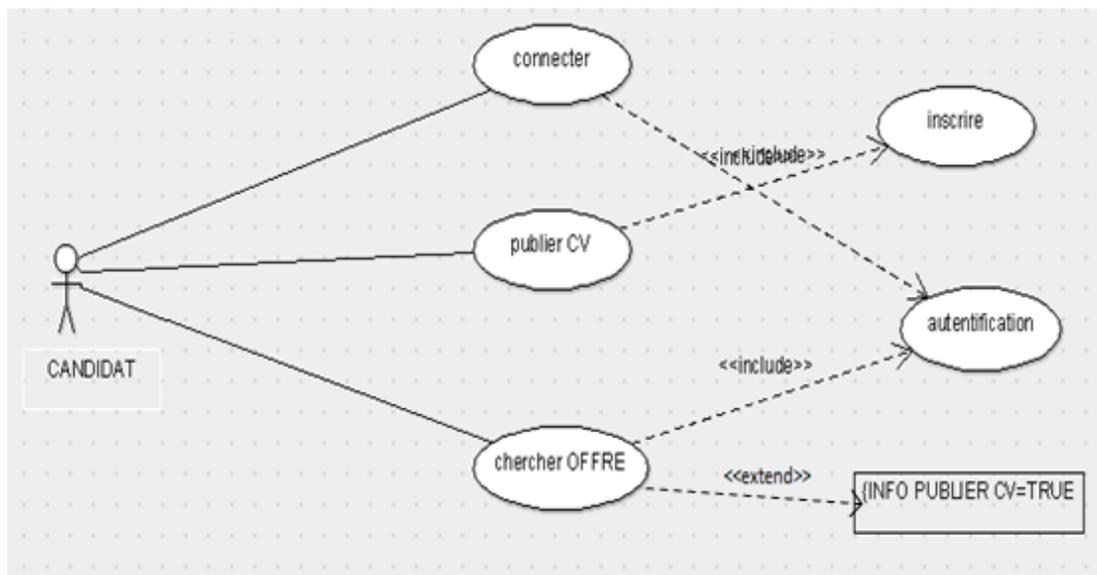


Figure III.7 – diagrammes de cas d'utilisation de candidat

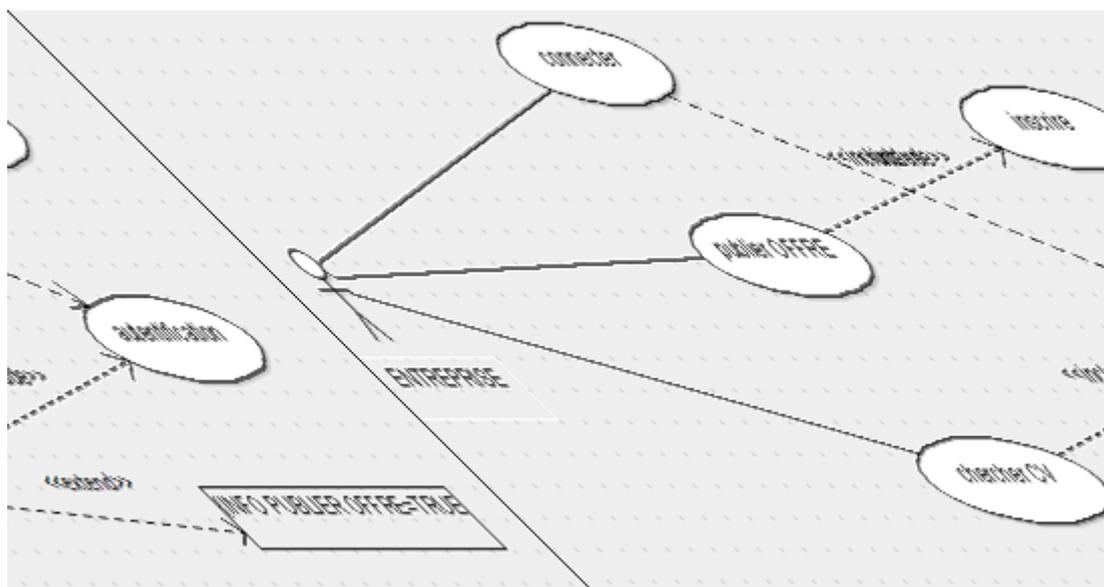


Figure III.8 – diagrammes de cas d'utilisation de l'entreprise

3. DIAGRAMME DE SEQUENCE : Les diagrammes de séquences permettent de représenter les interactions entre les objets en précisant la chronologie des échanges de messages. Ils peuvent être utilisés pour représenter les scénarios d'un cas d'utilisation donné.

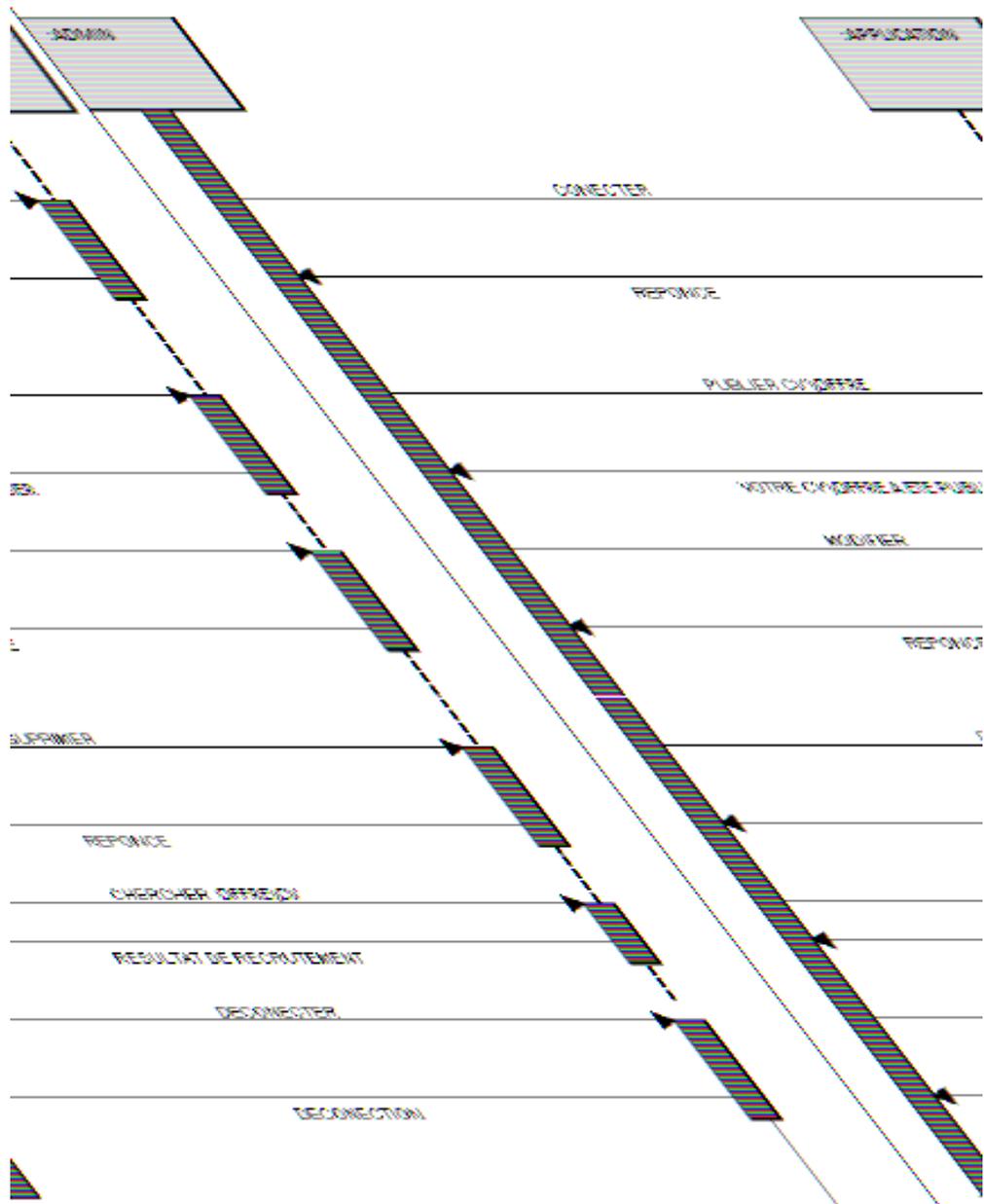


Figure III.9 – diagrammes de séquence de l'administrateur

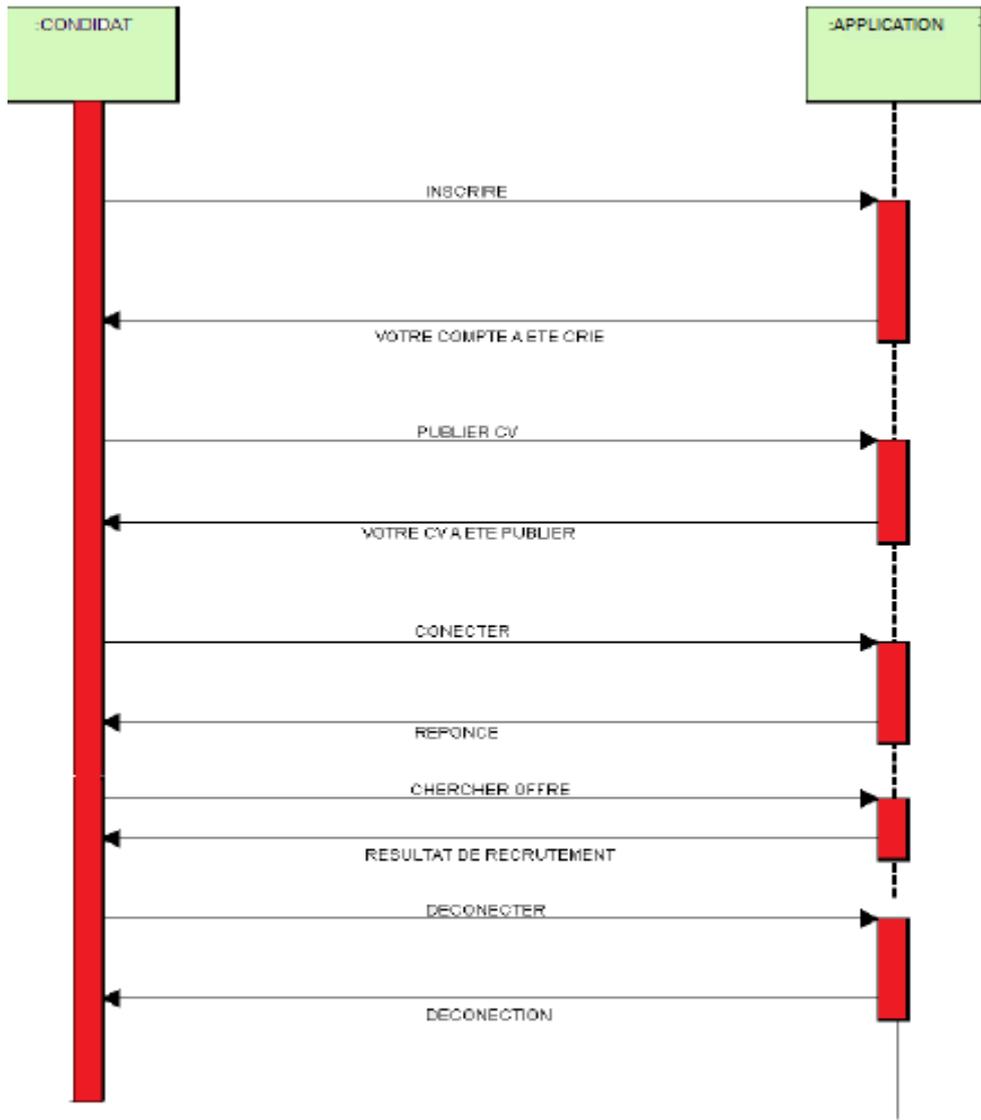


Figure III.10 – diagrammes de séquence de candidat

III.5.1 Les capture de l'application et l'explication de leurs fonctionnements

Nous allons montrer les principales fonctionnalités de l'application.

Figure III.12 c'est la page de connexion elle contient deux champs à remplir par l'administrateur.

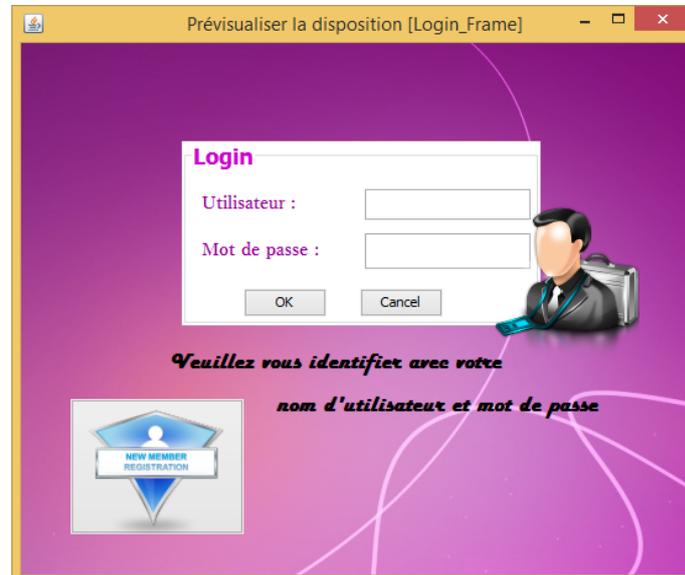


Figure III.12 –La page de connexion

La page suivante concerne la page d'accueil.

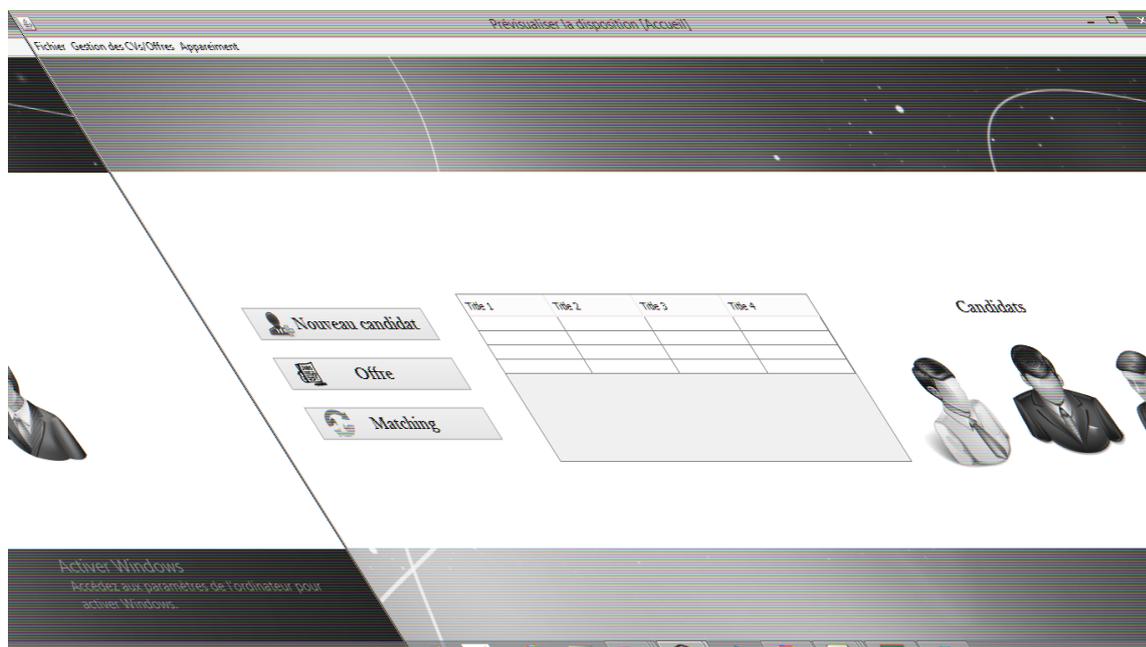


Figure III.13– La page d'accueil

C'est la page d'inscription qui va remplir le candidat, elle contient des informations qui va remplir le candidat, elle contient des informations banales concernant le nom, le prénom, l'âge ...etc.



The image shows a screenshot of a web browser window titled "Candidat". The page contains a registration form with the following fields and labels:

- Nom : [text input]
- Prénom : [text input]
- Sexe : [dropdown menu]
- Age : [dropdown menu]
- Tel : [text input]
- Mail : [text input]
- Adresse : [text input]
- Situation Famille : [dropdown menu]
- Conjoint : [text input]
- Nombre d'enfant : [text input]

On the left side of the form, there is a circular icon with a white exclamation mark on a dark background. On the right side, there is a stylized illustration of a man in a suit. At the bottom left, there is a watermark that reads "Activer Windows Accédez aux paramètres de Windows pour activer Windows."

Figure III.14 – La page d'inscription

La gestion des cv, c'est une fiche qui comprend des informations individuelles du candidat : le numéro du téléphone, l'adresse, situation familiale... etc.

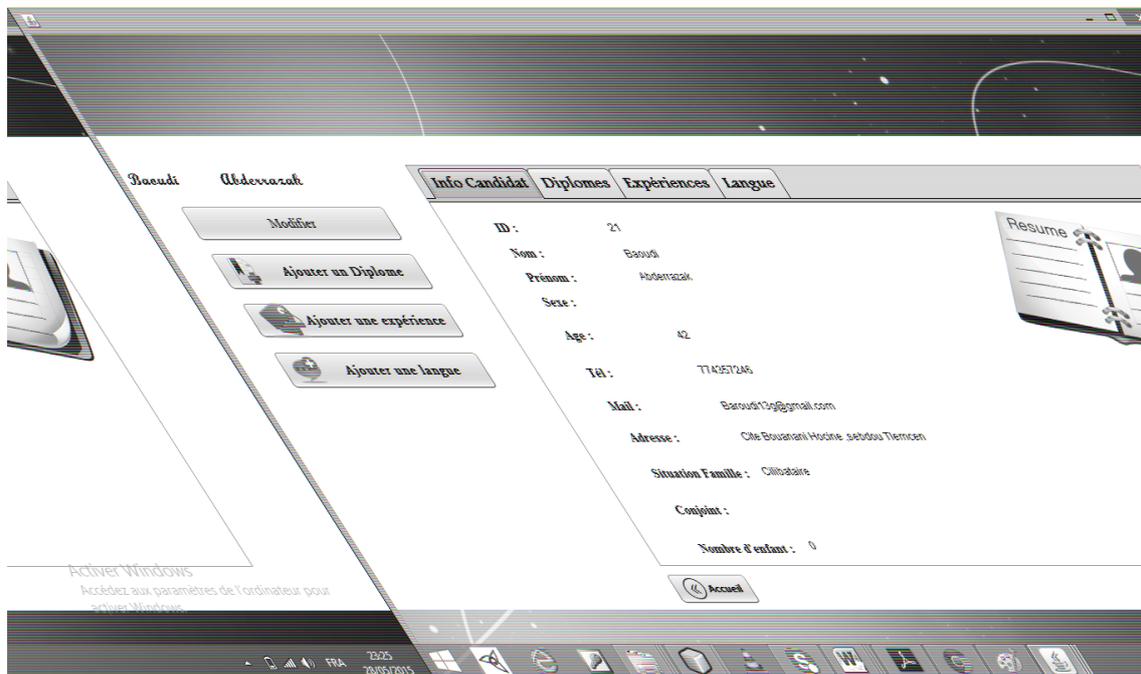


Figure III.15 – La gestion des cv

Cette figure montre la page d'ajout d'un diplôme, elle donne des informations concernant l'établissement, la durée et l'intitulé du diplôme.

Candidat :

Date début :

Date fin :

Intitulé diplôme :

Etablissement :

Ajouter

ActiveM

Figure III.16 – L’ajout d’un diplôme

La **Figure III.17** c’est celle de l’ajout d’une expérience du candidat, elle comprend : le type, la durée et la société ou c’est effectué cette expérience.

The image shows a web browser window displaying a form for adding a candidate's experience. The form is titled "Ajouter" (Add) and contains the following fields:

- Candidat :
- Type :
- Date début :
- Date fin :
- Société :
- Exp Fonction :
- Exp Salaire :

At the bottom left of the form is a button labeled "Ajouter" with a plus sign icon. The background of the form features a 3D figure holding a tray with a candle, and the word "EXPERIENCE" is visible on the left and right sides.

Figure III.17 – L'ajout d'une expérience

Figure III.18 c'est la page d'ajout d'une langue qui maîtrise le candidat.

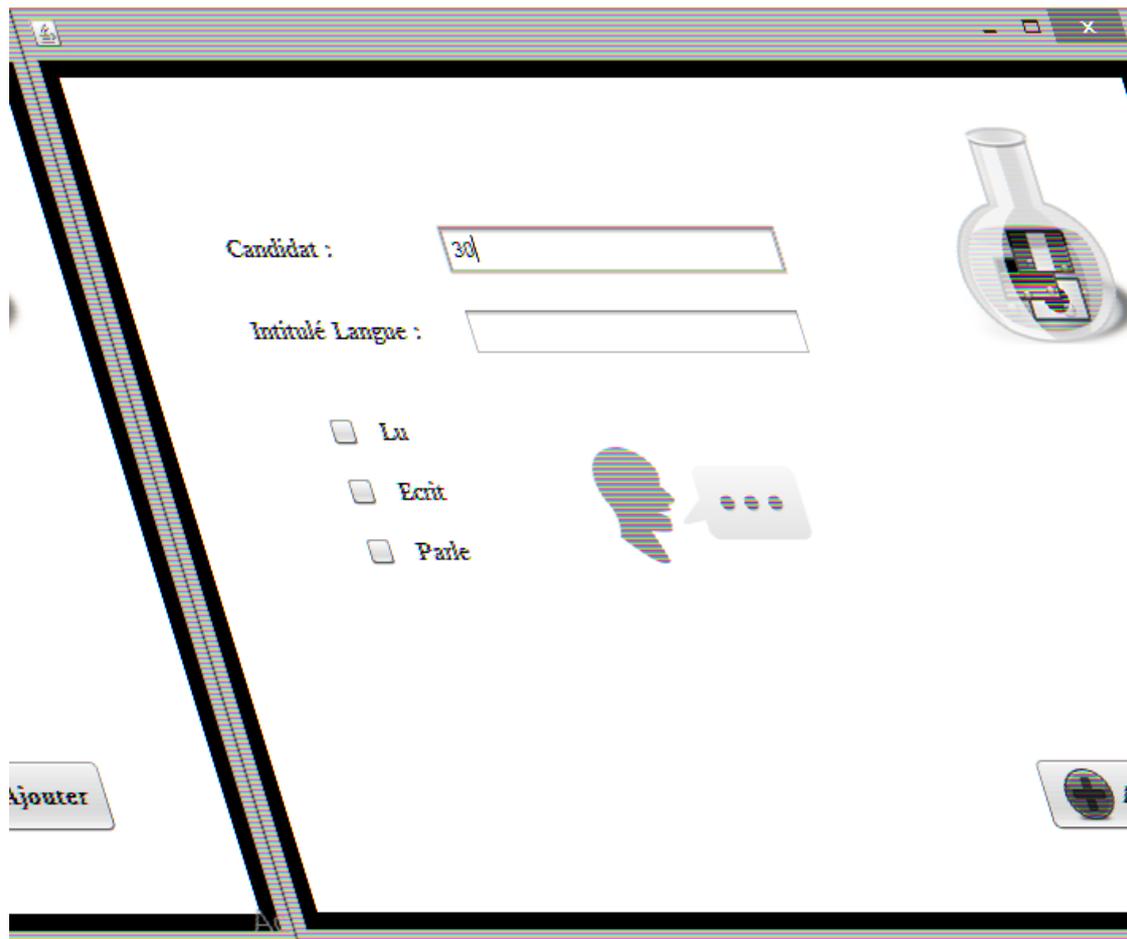


Figure III.18 – L’ajout d’une langue

Figure III.19 c’est la fiche concernant l’entreprise, c’est une page de l’offre d’emploi. Elle contient des informations concernant l’entreprise, le mode, le lieu, la durée, le salaire de cet offre d’emploi qui facilite au candidat de faire les premier démarches du contrat.

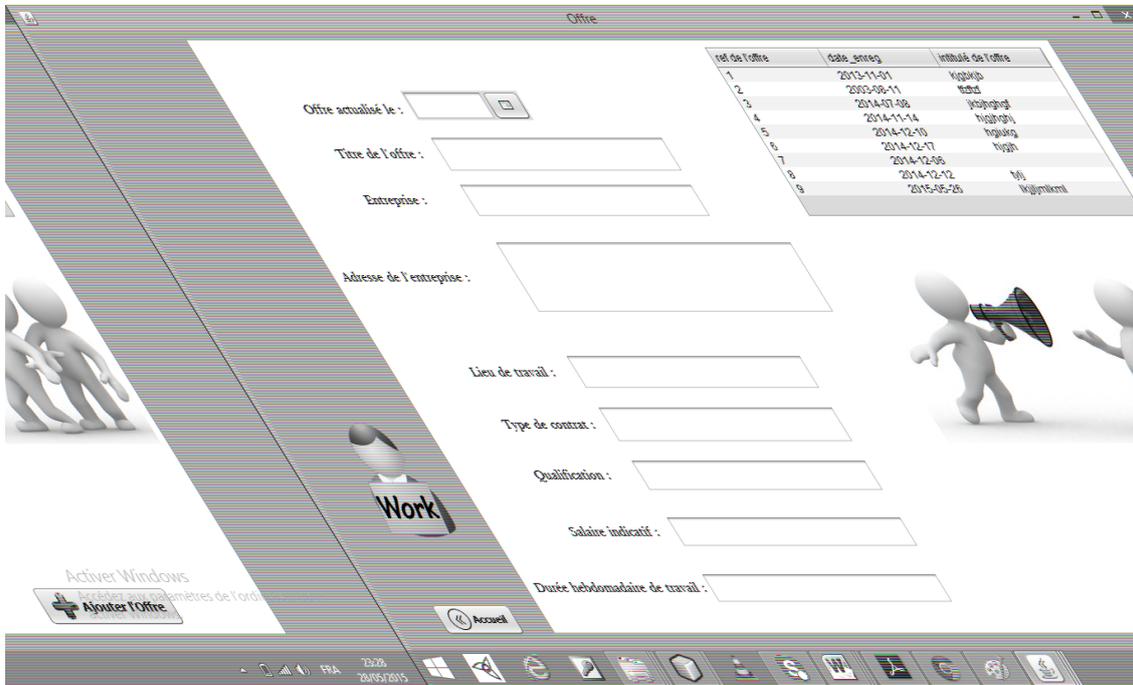


Figure III.19 – La gestion des offres

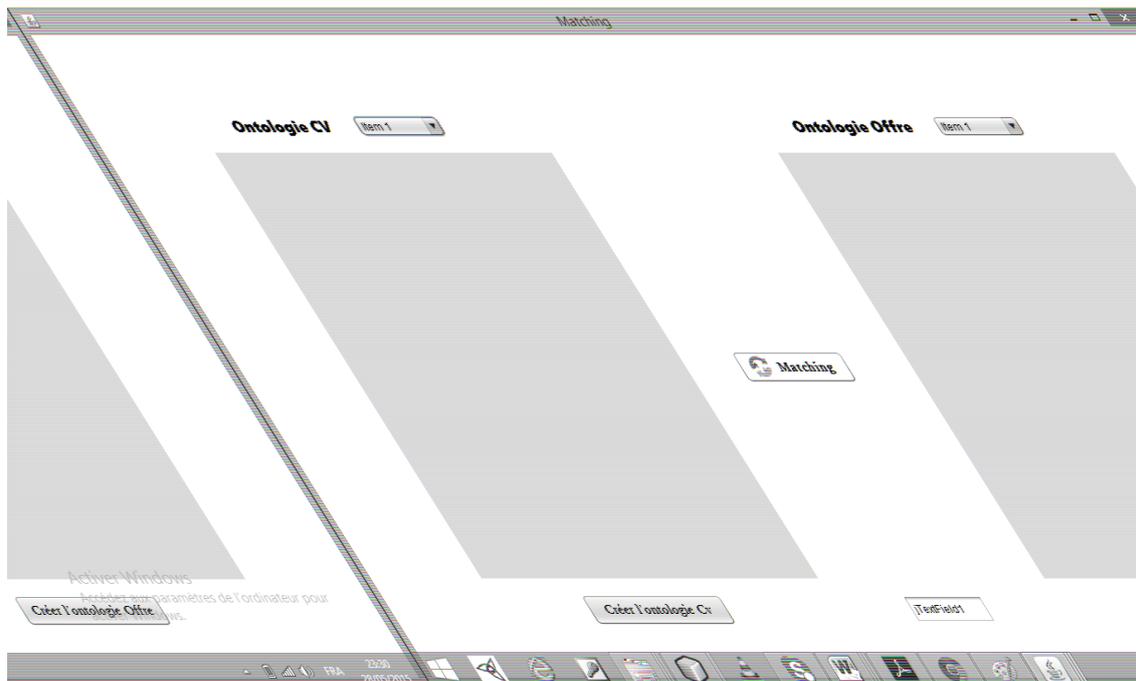


Figure III.20 – Page de Matching

III.6 CONCLUSION

Dans ce chapitre, on a proposé une modélisation du contenu sémantique à base d'ontologie des CVs et des offres d'emploi. Mais l'application de cette méthode n'a pas réussi, donc on a utilisé une autre méthode qui peut traiter facilement les CVs et les offres d'emploi.

CONCLUSION ET PERSPECTIVES

Conclusion :

L'objectif de ce mémoire était la conception et la réalisation d'un outil informatique, ayant pour but de faciliter le traitement des CVs et des offres d'emploi.

Dans ce travail, on a proposé une modélisation du contenu sémantique à base d'ontologie des CVs et des offres d'emploi. Mais l'application de cette méthode n'a pas réussi.

Donc on a proposé une autre méthode basé sur la comparaison mot à mot des CVs et des offres d'emploi.

Comme perspective on propose d'une part de rendre l'application sous forme une application web, et d'autre part d'annoter les textes par des ontologies.

Bibliographie

- [1] Arpirez J., Gómez-Perez A., Lozano A. et Pinto S., (ONTO)2Agent : An ontology-based WWW broker to select ontologies. Paper presented at the Workshop on Applications of Ontologies and PSMs, Brighton, England 1998.
- [2] Bachimont B. , Engagement sémantique et engagement ontologique : conception et réalisation d'ontologies en ingénierie des connaissances. In Z. M. Charlet J., Kassel G., Bourgault D., (Ed.), Ingénierie des connaissances. Évolution Récentes et nouveaux défis Paris : Eyrolles, 2000.
- [3] B. Bachimont, R. Troncy, A. Isaac., Semantic commitment for designing ontologies: a proposal. In Actes de la conférence "EKAW ", Siquenza, Spain. Springer Verlag .pdf,2000.
- [4] Bechhofer, S., Horrocks, I., Goble, C. et Stevens, R., OilEd : a Reason-able Ontology Editor for the Semantic Web. Paper presented at the Proceedings of KI2001, Joint German/Austrian conference on Artificial Intelligence" Vienna. Springer-Verlag, OilED : a Reasonable Ontology Editor the Semantic Web Proceeding of KI 2001.
- [5] Borst W. N, Construction of Engineering Ontologies, Center for Telematica and Information Technology, University of Twente, Enschede, NL1997.
- [6] DIENG R,CORBY O,GANDON F,GIBOIN A,GOLEBIEWSKA J,MATTA
- [7] Enrico Motta, Simon Buckingham Shum, and John Domingue. Ontology-driven document enrichment: principles, tools and applications, Knowledge Media Institute, janvier 2000.
- [8] Farquhar. A., Fikes. R, Pratt. W and Rice. J, Collaborative Ontologie Construction for Information Integrations, Technical Report KSL-95-63, Knowledge Systems Laboratory, Stanford University, 1995.
- [9] Fürst F., L'ingénierie Ontologique (02-07). Nantes : Institut de Recherche en Informatique de Nantes, 2002.
- [10] Gómez-Pérez A. (1999). Ontological Engineering : A state of the art. Expert Update, 2(3), 33-43.
- [11] GRUBER T,A translation approach to portable ontology specification ,knowledge Acquisition 5(2),pp199-220,1993.
- [12] Inaba A., Supnithi T., Ikeda M., Mizoguchi R. et Toyoda J, An Overview of Learning Goal Ontology, Proc. of ECAI2000 Workshop on Analysis and Modelling of Collaborative Learning Interactions, 2000.

- [13] [12] Jin L., Chen W., Hayashi Y., Ikeda M., Mizoguchi R., Takaoka Y. et Ohta M, An Ontology-Aware Authoring Tool - Functional structure and guidance generation. Paper presented at the AIED1999, Le Mans, France.1999.
- [14] Maedche A, Ontology Learning for the Semantic Web. Boston : Kluwer Academic Publishers,2000.
- [15] Mizoguchi R., Kozaki K., Sano T. et Kitamura Y. , Construction and Deployment of a Plant Ontology. The 12th International Conference, EKAW2000.
- [16] Mizoguchi R, A Step Towards Ontological Engineering. Paper presented at the 12th National Conference on AI of JSAI, June, 1998.
- [17] N.GUARINO ,Understanding building and using ontologies.Acommentary to using explicit ontologies in KBS development, by van heijst,Schreiber,and wielinga,1997.
- [18] Rose Dieng , Olivier Corby , Fabien Gandon , Alain Giboin , Joanna Golebiowska , Nada Matta, Myriam Ribiere , méthodes et outils pour la gestion des connaissances :une approche pluridisciplinaire du knowledge management ,dunod Edition informatique Séries systèmes d'information 2001.
- [19] Stab s,H-P Schnurr,R Studer ,Y Sure, knowledge process and ontologies, IEEE intelligent System,January/February 2001,val,16,num 1 special Issue on Knowledge Management.
- [20] Sowa J, Distinction, combination and constraints, Proc. IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, (1995a).
- [21] Studer R., Benjamins R. et Fensel D. , Knowledge Engineering : Principles and Methods. Data Knowledge Engineering, 1998.
- [22] Sure et al, OilED : a Resonable Ontology Editor the Semantic Web Proceeding of KI 2001.
- [23] Uschold M. et King M., Towards a Methodology for Building Ontologies, Paper presented at the Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.
- [24] Van Heijst G., Schreiber A. et Wielinga B. J, Using Explicit Ontologies in KBS Development. International Journal of Human and Computer Studies /Knowledge Acquisition, 46(2/3), 183-292, . (1997).
- [25] Vanwelkenhuysen J. et Mizoguchi R,Maintaining the workplace context in a knowledge level analysis,. Paper presented at the Proc. of JKAW'94, Hatoyama, Japan, . (1994).
- [26] Wielinga B. et Schreiber A. , Reusable and sharable knowledge bases : A European perspective, Paper presented at the KB and KS 93, Tokyo, (1993).
- [27] Fouzia Amourach, L.Yahiaoui, construction d'une ontologie pour l'annotation des CVS/offres d'emploi, université de mentouri, Constantine 2008.
- [28] avantage du web sémantique, HowPublished = <http://solutions.journaldunet.com/printer/031003-tribune.shtml>.

[29] Charlet J., Laublet P; Reynaud C, Web Sémantique : Rapport final de l'action spécifique, 32 du CNRS/STIC, Dec. 2 <http://rtp-doc.enssib.fr/IMG/pdf/ASWebSemantique2003.pdf>

[30] GHONIEM Mahmoud , NIKIEMA Joseph, Rapport de projet Sur Le Web sémantique, howPublished = <http://solutions.journaldunet.com/printer/031003-tribune.shtml>,

[31] Luis Adrián Cabrera-Diego, Juan-Manuel Torres-Moreno, Marc El-Bèze, Seg CV : traitement efficace de CV avec analyse et correction d'erreurs, TALN-RÉCITAL 2013.

[32] Sabri Boutemedjet, le Web sémantique

[33] Zouhel Boucetta , Zizette Boufaïda , Leïla Yahiaoui, l'appariement sémantique des documents à base d'ontologie pour le e-Recrutement, Laboratoire LIRE - Université Mentouri Constantine- Algérie.

[34] Travail d'intérêt personnel encadré RAPPORT FINAL Ontologies et Web Services Réalisé par PHAN Quang Trung Tien Promotion X Sous la direction de Professeur NGUYEN Hong Quang.

Site d'internet :

[35] Amandine Schuurman Community web portails références, projet de Master , how Published =: <http://www.ics.forth.gr/christop/portails.htm>

[36] Bertrand Sajus, La fonction Thésaurale au cœur des systèmes d'information, ADBS, HowPublished =www.adbs.fr/adbs/prodserv/jetude/html/prog110402a.html, avril,2002

[37] Charlet J., Laublet P; Reynaud C, Web Sémantique : Rapport final de l'action spécifique 32 du CNRS/STIC, Dec. 2 <http://rtp-doc.enssib.fr/IMG/pdf/ASWebSemantique2003.pdf>

[38] cycle de vie des ontologies , HowPublished=<http://intercities.info/jcms/c-17672/ontologies-informatique>.

[39] jdbc, HowPublished =http://www.commentcamarche.net/content/java/java_jdbc.php3

[40] Netbeans, HowPublished =<http://www.techno-science.net/?onglet=glossaire&definition=5346>

[41] Recommandation OWL sur le site du W3C, How Published =<http://www.w3.org/2004/OWL/>

[42] Recommandation RDF sur le site du W3C, How Published =<http://www.w3.org/RDF/>

[43] Site Web du projet Web sémantique de HP (Jena), HowPublished = <http://www.jena.sourceforge.net>

[44] source en français sur XML, How Published = <http://xmlfr.org/>

[45] Tim Berners-Lee, James Hendler, and Ora Lassila., The Semantic Web, howPublished=www.scientificamerican.com/article.cfm?articleID=00048144-10D2-1C7084A9809EC588EF21&catID=2, May 2001.

[46] TuanAnh Ta, Web sémantique et portails un état de l'art, HowPublished =www.infres.enst.fr/people/saglio/etudes/e-parcours/portails/SemWebintro.pdf,2001.

Liste des figures

Figure I.1 – architecture du web sémantique	15
Figure II.1 – le cycle de vie de l'ontologie.....	25
Figure II.2 – typologie de l'ontologie selon les dimensions de classification.....	27
Figure II.3 –typologie selon l'objet de conceptualisation.....	31
Figure II.4 – typologie selon le niveau de détail.....	32
Figure II.5 – typologie selon le niveau de complétude.....	32
Figure II.6 – typologie selon le niveau du formalisme.....	33
Figure III .1 architecture du système.....	39
Figure III .2 – Editeur NetBeans.....	40
Figure III.3 – diagrammes d'activité d'administrateur.....	42
Figure III.4 – diagrammes d'activité de candidat.....	42
Figure III.5 – diagrammes d'activité de l'entreprise.....	43
Figure III.6 – diagrammes de cas d'utilisation de l'administrateur.....	44
Figure III.7 – diagrammes de cas d'utilisation de candidat	44
Figure III.8 – diagrammes de cas d'utilisation de l'entreprise.....	45
Figure III.9 – diagrammes de séquence de l'administrateur.....	46
Figure III.10 – diagrammes de séquence de candidat.....	47
Figure III.11 – diagrammes de séquence de l'entreprise.....	48
Figure III.12 –La page de connexion	49
Figure III.13– La page d'accueil.....	49
Figure III.14 – La page d'inscription.....	50
Figure III.15 – La gestion des cv.....	50
Figure III.16 – L'ajout d'un diplôme	51
Figure III.17 – L'ajout d'une expérience.....	51
Figure III.18 – L'ajout d'une langue	52
Figure III.19 – La gestion des offres	53
Figure III.20 – Page de Matching	53

Liste des Tableaux

Tab I.1 – différence entre le web traditionnels et le web sémantique.....7

Résumé:

Le marché d'offres d'emploi et des candidatures sur Internet ont connu, ces derniers temps, une croissance exponentielle. Ceci implique des volumes d'information intraitables manuellement. Les CV sont dans des formats très divers : formulaire, .pdf, .doc, .dvi, .ps, etc., ce qui entraîne une mauvaise recherche.

Dans ce travail, nous proposons des ontologies pour la modélisation du contenu sémantique des CVs/Offres d'emploi dans le cadre d'automatisation de la recherche d'emploi.

Les ontologies construites se concentrent sur une modélisation explicite des informations contenues dans les CVs/Offres. Nous appliquons les mesures de similarité sur ces ontologies afin d'assurer une automatisation de recherche d'un offre d'emploi.

Mots clés : *Web sémantique, ontologie, CVs/Offres d'emploi.*

Abstract :

Over the last years, the online market of jobs and candidatures offers has reached an exponential growth. This has implied great amounts of information which cannot be processed manually. The CV are in several formats: formular, .pdf, .doc, .dvi, .ps, etc., that can provoke errors in the search.

In this work, we propose an ontology for the modelisation of the semantic content in terms of their supply and demand, so it allows their automatic matching.

Due to the importance of the underling competencies of these documents, for their matching. The building ontology is focused on the explicit modelisation of the competency. The adopted model assures a certain compromise between the simplicity and the pertinence, distinguishing the identification of acquired demanded competencies and explicating the semantic of the relation can be exists between the different topics of competencies.

Key Words: *semantic Web, ontology, CV/Job Offer.*

صخلم

فقبح وعصنلوي امملمعلكابلطاذلكي نرتن ألكبش لعلمعلاض ورعثي ح نم له ندمعاصت دوج وظحولة تري خالته اوان سنس لي في الكش أ تدعي فضرع هوي تاذقل يسل ان أ وقص لخي حشرتمل ا فرط نم دم دق مله ا فل م اة ج ل ا عم

راقملا رنك م ت ل ل لم عل ك ا ب ل ط و ض و ر ع ل ي و ن ع م ل ف ا و ت ح م ل ل ج ذ م ن ا ط ا ع ل ج ا ل ي م ج و ل و ت ن ل م ع ل ا ا ذ ي ف ع ر ت ق ن ا ث و ل ا ه ن ق ي ب ك ي ت ا م و ت و أ ل ا

لت و ه ج ق ي م س ت ك م ل ت ا ر د ق ل ن ا ي ب ق ر ف ل ا ب ت ع ا ل ا ي ع ب ذ خ أ ل ا عم ، د ح ا و ن ا ي ف ا ة ي م ه أ ل ا و ق ل و ه س ل ن ا ي ق ي ف ا و ت ي م ض ج ر ت ق م ل ا ج ذ م ن ل ا ن ا ت ا ر د ق ل ع ي ض ا و ف ل ت خ ف ي ب ق م ا ل ق ل ق ا ل ع ل ن ا ي ح ي ض و ت و ، ي ر خ أ ه ج ن ق ب و ل ط م ل ا

لمع ن ع ج ب ل ة ي ل م ع ه س ي ل م م ا ن ا ي ب ل ل ل م ع ت س ا ب ح ض ل و ك ش ه ب ق ي ن ق ي ذ غ ت ل ب م ع ت س م ل ا ج م ب ل ي ب ح ل و ل و ت ن أ ل ا ه ذ ه ، ي ر خ أ ه ج ن م

لمعلكابلط و ض و ر ا ج و ل و ت ن ا ي و ن ع م ل ا ب ا ق ل ي ا ح ا ت ف م ل ه ا م ل ك ل ا