

**UNIVERSITE KASDI MERBAH OUARGLA**

**Faculté des Nouvelles Technologies de l'Information et de la Communication**

**Département d'Informatique et Technologie de l'information**



**Mémoire de fin d'études**

**Pour l'obtention du diplôme de Master en Informatique**

Domaine : Mathématique et Informatique

Filière : Informatique

Spécialité : Informatique Industrielle

Présenté par : **HADEF Mourad**

**MEHAOUA Fares**

**Thème :**

**Prise en compte du profil utilisateur  
Dans un système de recherche d'information**

Soutenu publiquement

Le : 12 Juin 2014

*Devant le jury composé de :*

**M. DJEDIAI Hmida**

*(Président)*

**UKM OUARGLA**

**M. ZGA Adel**

*(Membre)*

**UKM OUARGLA**

**M. BELOUAAR Houcine**

*(Reporteur)*

**UKM OUARGLA**

Année universitaire: 2013-2014

## **Remerciements**

*HADEF Mourad et MEHAOUA Fares aimeraient avant tout remercier l'unique « Allah » pour tous ses bienfaits trop souvent négligés.*

*Tout au long de ce projet, nous avons été très bien entourées par des personnes qui nous ont guidées, inspirées, encouragées, critiquées et surtout, qui nous ont permis d'avancer.*

*Nous prenons ici le temps de sincèrement les remercier et nous leur devons à tous, une partie importante de cette recherche.*

*Nous remercions également Mr Houcine BELOUAAR, Maitre-assistant au Département d'Informatique à l'université de Kasdi Merbah Ouargla pour avoir accepté de suivre notre projet de manière honnête et intègre et donc de lire consciencieusement ce mémoire. Vos nombreuses remarques, ont été la preuve de l'intérêt porté à notre travail, ce en quoi nous vous sommes extrêmement reconnaissantes.*

*Nos remerciements s'adressent aussi aux membres du Jury pour nous avoir honorées en consentant à juger notre modeste travail.*

*Notre reconnaissance va également à l'ensemble des enseignants de Département d'Informatique , qui nous ont permis, grâce à leurs savoirs et à leurs persévérances, d'en arriver jusqu'ici, qu'ils reçoivent tous nos sentiments de respects, de considérations et de reconnaissances.*

*Enfin, nous remercions toutes les personnes qui, scientifiquement ou socialement ont contribué à la réussite de ce travail.*

## *Dédicace*

*Nous dédions ce mémoire,  
à tous ceux et toutes celles  
qui nous ont accompagnés et soutenu  
durant cette année d'étude*

## Résumé

Le profil utilisateur est utilisé dans les logiciels collectifs, tels que les systèmes d'exploitation, les systèmes de gestion de bases de données, les moteurs de recherche ou les sites de e-commerce.

Dans les systèmes d'exploitation Windows, un profil utilisateur contient les réglages personnels appliqués à l'environnement de bureau pour cet utilisateur; par exemple les marque-page du navigateur web, la connexion au réseau, les réglages des imprimantes, les comptes e-mail ou les documents personnels.

Dans le logiciel de base de données Oracle les profils utilisateurs permettent de limiter la quantité de ressources matérielles. Chaque profil peut être attribué à plusieurs utilisateurs, et permet par exemple de limiter l'utilisation du processeur, la durée des sessions ou la longueur des mots de passe.

L'objet de notre travail est de modéliser le profil utilisateur pour ensuite être utilisé dans le domaine de recherche d'information et dans le but de mieux filtrer les informations, le profil sera présenté seulement par le domaine d'intérêt de l'utilisateur.

**Mot clés :** système de recherche d'information, profil utilisateur, modélisation du profil, ontologie.

## Abstract

The user profile is used in collaborative software , such as operating systems, management of databases, search engines or sites e -commerce systems.

In Windows operating systems , a user profile contains personal settings applied to the desktop environment for the user ; eg bookmark web browser, connect to the network settings of printers, e -mail accounts and personal documents.

In the software Oracle database user profiles can limit the amount of material resources. Each profile can be assigned to multiple users, and allows for example to limit CPU usage , duration of sessions or the length of passwords.

The purpose of our work is to model the user profile and then be useful in the field of information retrieval and to better filter information , the profile will be presented only in the domain of interest of user.

**Keyword :** system information retrieval, user profile, profile modeling , ontology.

# Table des matières

Introduction générale .....	1
Problématique .....	1
Organisation du mémoire.....	1

## Chapitre I : les systèmes de recherche d'information

I.1.Introduction .....	2
I.2.Définition d'un Système de Recherche d'Information .....	3
I.3.Besoins d'information .....	4
I.4.Caractéristiques du besoin d'information .....	5
I.5. Les concepts fondamentaux de la recherche d'information .....	5
I.6.Les types de systèmes de recherche d'information .....	7
I.6.1.Le système de recherche d'information classique.....	7
I.6.2.Le système de recherche d'information adaptatifs.....	7
I.6.3.Le Système de Recherche d'Information personnalisé .....	7
I.9.Défis d'un SRI .....	8
I.10.Conclusion .....	9

## Chapitre II: le profil utilisateur

II.1. Introduction.....	10
II.2. Notion de profil.....	11
II.2.1. Le profil utilisateur .....	11
II.2.2. Le profil document.....	11
II.3. Processus de définition du profil de l'utilisateur .....	12
II.3.1. Représentation.....	12
II.3.2. Construction .....	13
II.4. Le contenu du profil.....	13
II.4.1. Domaine d'intérêt.....	14
II.4.2. Données personnelles .....	14
II.4.3. Qualité .....	15
II.4.4. Données de livraison .....	16
II.4.5. Données de sécurité .....	17
II.5. Exemple d'un système de recherche d'information utilisant le prof : STRIP .....	18
II.5.1. Modélisation du profil utilisateur .....	20

II.5.2. Exploitation du profil d'utilisateur.....	21
II.6. Conclusion .....	21

## Chapitre III: la modélisation du profil utilisateur

III.1 Introduction .....	22
III.2 Le filtrage d'information collaboratif pour de modélisation du profil.....	23
III.2.1 Filtrage de l'information.....	23
III.2.2 Modélisation d'un profil .....	25
III.3 Techniques statistiques pour la définition du profil d'un utilisateur.....	27
III.4 Techniques base sur l'ontologie .....	29
III.4.1 Notion d'ontologie.....	29
III.4.2 Cycle de vie d'une ontologie .....	31
III.4.3 Principes de développement d'une ontologie .....	31
III.5 Les composants d'une ontologie .....	32
III.6 Méthodologie de l'ingénierie ontologique .....	33
III.7 Les types d'ontologies .....	34
III.8 Modélisation du profil utilisateur base sur l'ontologie .....	34
III.9 Conclusion.....	38

## Chapitre IV : conception et réalisation

### Partie 1: la conception

IV .1 .Introduction.....	39
IV.2.UML .....	40
IV.2.1. Les Diagrammes UML .....	40
IV.2.1.1. Diagrammes structurels ou statiques.....	40
IV.2.1.2. Diagrammes comportementaux.....	40
IV.2.1.3. Diagrammes d'interaction ou dynamiques.....	40
IV.3. Un réseau social .....	41
IV.4. Conception .....	41
IV.4.1. Conception l'ontologie.....	41
IV.4.2.Architecture générale du système .....	46
IV.5.Description du système.....	47
IV.5.1.Module d'inscription .....	47
IV.5.2.Module de recherche des amis.....	47
IV.5.3.Module d'interaction avec les amis .....	47

IV.5.4. Module de configuration .....	47
IV.6. Le Diagramme de cas d'utilisation .....	48
IV.7. Les diagrammes de séquences.....	51
IV.8. Conclusion .....	54
<b>Partie 2: la réalisation</b>	
IV.9. Introduction.....	55
IV.10. Environnement de développement .....	55
IV.10.1 Protégé .....	55
IV.10.2. Le langage SWRL .....	55
IV.10.3. Le langage Java.....	56
IV.10.4. NetBeans .....	56
IV.11. Scénarios d'exécution avec captures d'écrans .....	57
IV.12. Conclusion .....	64
<i>Conclusion générale et perspective.....</i>	<b>65</b>

## Bibliographie

## Liste des figures

---

Figure I.1 : Processus en U de la RI .....	6
Figure I.2 : Système de recherche d'information .....	8
Figure II.1: Méta modèle de la dimension domaine d'intérêt .....	14
Figure II.2 : Méta modèle de la dimension donnée personnelles .....	15
Figure II.3: Méta modèle de la dimension Qualité .....	16
Figure II.4: Méta modèle de la dimension donnée de livraison .....	17
Figure II.5 : Méta modèle de la dimension sécurité .....	18
Figure II.6: Processus global d'accès personnalisé à l'information .....	19
Figure III.1: Filtrage basé sur le contenu.....	24
Figure III.2: Processus de filtrage collaboratif .....	24
Figure III.3: Modélisation d'un profil utilisateur .....	26
Figure III.4: Différentes acceptations du terme ontologie.....	30
Figure III.5 : schéma de l'ontologie .....	30
Figure III.6 cycle de vie d'une ontologie .....	31
Figure III.7: Processus de construction d'ontologie.....	33

Figure III.8:Modélisation du profil utilisateur.....	34
Figure III.9: ontologie apprenant élaboré .....	35
Figure III.10 : L'ontologie utilisateur (concepts abstraits)-en OI Modeler.....	36
Figure III.11 : la dimension donnée de livraison.....	37
Figure IV.1 : Le diagramme de classe de l'ontologie .....	45
Figure IV.2: Le Diagramme de classe profile danse application .....	46
Figure IV. 3: L'architecture générale du système .....	46
Figure IV.4 : Diagramme de cas d'utilisation de utilisateur.....	48
Figure IV.5: Diagramme de cas d'utilisation de Système .....	49
Figure IV.6 : Le diagramme de séquence pour l'authentification.....	51
Figure IV.7 : Le diagramme de séquence pour la recherche .....	52
Figure IV.8 : Le diagramme de séquence pour la modification .....	53
Figure IV.9 : l'interface d'application.....	57
Figure IV.10 : le test de CAPTCHA.....	57
Figure IV.11 : Présentation des informations personnelles .....	58
Figure IV.12:créer un groupe .....	58
Figure IV.13: la page Home .....	59
Figure IV.14: le profil.....	59
Figure IV.15: information personnelle .....	60
Figure IV.16: Les personnes trouvées .....	60
Figure IV.17: Ajouter la photo de couverture et de profil.....	61
Figure IV.18: la page groupe.....	62
Figure IV.19: « Room chat » dans un groupe.....	62
Figure IV.20: interaction par Message .....	63

## Liste des tableaux

---

Tableau IV.1 : Les classes et hiérarchie des classes de l'ontologie. ....	42
Tableau IV.2 : Les propriétés du modèle de l'ontologie.....	42
Tableau IV.3 : Les Relations entre concepts du modèle de l'ontologie .....	43
Tableau IV.4 : Des exemples d'instances des classes du modèle d'ontologie. ....	44
Tableau IV.5 : Identification des scénarios .....	50

## **Introduction générale**

De nos jours, nous sommes privilégiés dans un monde riche en information, dans lequel la plupart, si ce n'est la totalité, de l'information dont nous avons besoin est au bout de nos doigts et est prête à être exploitée. L'essor du Web, conforté par le rapide développement des nouvelles technologies de l'information et de la communication a conduit à la production d'un volume d'information sans précédent, et la quantité d'information manipulée par diverses organisations dans le monde est élevée et sa gestion sans l'ordinateur n'est plus imaginable. Ce flux informationnel est ingérable avec les moyens classiques, ainsi le moment s'avère opportun pour l'emploi de techniques permettant une gestion lucrative de la connaissance humaine.

## **Problématique**

La difficulté d'évaluer l'information désirée a donc augmenté avec la croissance du volume de l'information disponible, ce qui engendre un manque d'outils suffisants pour aller vers une exploitation maximale de l'information disponible et éliminer le problème de surcharge.

Ces contributions soulèvent à la communauté de la recherche d'information (RI) de nouveaux défis, d'énormes efforts ont été déployés ,pour développer des approches et des techniques permettant de retrouver l'information voulue effectivement et efficacement à partir de vastes collections de données. Notamment avec l'avènement d'Internet, la recherche d'information est devenue plus d'actualité et plus exigeante que jamais.

## **Organisation du mémoire**

Notre mémoire est organisé comme suite :

Le premier chapitre est consacré aux concepts fondamentaux de la recherche d'information et les défis d'un système de recherche d'information.

Dans le second chapitre, nous abordons la notion du profil utilisateur et le processus de définition du profil utilisateur et nous terminons ce chapitre par un exemple d'un système de recherche d'information SRIP.

Le troisième chapitre présente les techniques utilisée pour la définition du profil, la notion d'ontologie et son intérêt dans la modélisation du profil, puis nous présentons notre modélisation basé sur l'ontologie.

Enfin, dans le quatrième chapitre, nous présentons la conception de notre solution (les différents graphes UML) et nous terminons ce chapitre par une implémentation qui a comme objet la réalisation d'un chat à base du profil utilisateur.

## Chapitre I

# Les Systèmes de recherche d'information

## I.1. Introduction

Les récentes avancées des technologies de la communication ouvrent de nouvelles perspectives. En particulier, les autoroutes de l'information permettent une diffusion instantanée d'une grande quantité d'informations.

Face à ce déploiement technologique il convient de s'interroger sur la qualité de l'information et sa maîtrise intellectuelle. Dans ce contexte le problème n'est plus tant la disponibilité de l'information mais la capacité de sélection d'une information qui réponde au besoin d'un utilisateur, à partir des représentations qu'il perçoit.

En fait, cette information est disponible mais souvent inaccessible. Plus que jamais, il faut donc assister l'utilisateur pour retrouver l'information dont il a besoin. Il est par exemple humainement difficile de parcourir toutes les pages du WWW pour rechercher une information particulière.

La mise en place de moteurs de recherche ne résout pas les problèmes de l'utilisateur qui obtient souvent plusieurs centaines de réponses à ses requêtes, présentées sous la forme d'une liste.

Les systèmes de recherche d'information (SRI) à concevoir doivent permettre l'accès à une grande quantité d'information par tout public en prenant en compte la variété des environnements d'utilisation à domicile [1].

## I.2. Définition d'un Système de Recherche d'Information

La recherche d'information (RI) est un domaine de l'informatique qui s'intéresse à la représentation, au stockage, à l'organisation et à la sélection d'information répondant aux besoins des utilisateurs.

Un système de recherche d'information peut être amélioré en, modélisant, intégrant, exploitant le contexte. Ainsi, le contexte peut être utilisé par exemple pour améliorer la façon dont les individus formulent leurs besoins au sein du système de recherche d'information et explorent les informations.

Traditionnellement, des variables contextuelles importantes sont incluses : les contextes d'utilisateur (par exemple, ses centres d'intérêts à court et long terme, ses habitudes...) ; les contextes d'objet ; les tâches et les contextes sociaux dans lesquels les besoins informationnels surgissent [2].

Il existe plusieurs types de dimensions pouvant être intégrés pour la contextualisation des systèmes de recherche d'information. Nous pouvons citer des approches qui permettent d'enrichir ou de raffiner la requête initiale des usagers.

Si nous prenons l'exemple du domaine d'intérêt de l'utilisateur, le résultat de recherche du terme « Python » ne doit pas être le même pour un informaticien et un vétérinaire. En intégrant le domaine de l'utilisateur le système de recherche d'information pourra ajouter les termes « langage de programmation » à la requête initiale de l'informaticien [2].

Les systèmes de recherche d'information contextualisés peuvent prendre en compte d'autres aspects du contexte, comme la nature de la tâche ou l'environnement de la recherche (localisation, matériel disponible, etc.), pour adapter le processus de recherche.

Les chercheurs dans le domaine système de recherche d'information Cool et Spink distinguent quatre niveaux pour le contexte dans la recherche d'information : le niveau environnemental de l'information, le niveau de recherche d'information, le niveau interactionnel de la recherche d'information et le niveau de la requête.

La plupart des travaux s'accordent sur un cœur commun qui inclut l'environnement et les dimensions humaines, mais divergent sur les éléments qui doivent être inclus dans le contexte.

À l'heure actuelle, presque toute la recherche dans ce secteur est dépendante de la spécification explicite des buts de recherche, des tâches et des intentions de l'usage [2].

### **I.3. Besoins d'information**

La prise en compte de l'utilisateur et de son besoin est souvent sous-estimée dans la conception et l'évaluation d'un système de recherche d'information. Les travaux dans le domaine des systèmes de recherche d'information s'intéressent à des modèles de recherche d'information qui sont capables de déterminer les documents qui répondent le mieux à une requête.

Un modèle de recherche d'information, est composé des représentations des documents, des représentations des requêtes d'un utilisateur, et de la fonction de correspondance entre ces deux éléments [3].

L'évaluation d'un système de recherche d'information se fonde sur des courbes de rappel/précision qui fournissent une mesure relative à une collection de documents et de requêtes résolues. Ces collections tests sont des "étalons de mesure" fabriqués par des experts.

Les résultats obtenus traduisent seulement l'aptitude moyenne du système sur des recherches en une seule requête l'interaction avec l'utilisateur et l'itération de la recherche sont totalement ignorées.

L'utilisateur, son besoin et son activité ne sont donc pas pris en compte dans le processus d'évaluation actuel de ces systèmes, ce qui est paradoxal pour des systèmes dont le but est justement de répondre à des demandes d'information de l'utilisateur [3].

## **I.4. Caractéristiques du besoin d'information**

Cet axe identifie des propriétés intrinsèques du besoin d'information en faisant abstraction de sa formulation en entrée du système. Ces propriétés sont conditionnées par l'état cognitif de l'utilisateur.

Les connaissances de l'utilisateur vis-à-vis du domaine de la recherche sont des facteurs primordiaux dans la formulation du besoin d'information car elles lui permettent de mieux cerner son besoin d'information.

Pour un domaine donné, chaque utilisateur a sa propre expertise, liée à ses connaissances et à son expérience, De son côté, le système a lui-même une connaissance du domaine sous la forme des index, du thesaurus, ou plus généralement sous la forme d'une base de connaissance [4].

Sur le domaine du corpus, les points de vue du système et de l'utilisateur ne sont donc pas forcément cohérents, à cause de la partialité ou de l'imprécision des connaissances que possède le système ou bien à cause du manque d'information que l'utilisateur peut avoir sur le domaine de la recherche.

Ainsi, un besoin d'information peut être bien défini dans l'esprit de l'utilisateur et complètement flou vis-à-vis du système. Le besoin d'information va alors se préciser au fur et à mesure que l'utilisateur va acquérir des connaissances grâce aux documents proposés par le système [4].

## **I.5. Les concepts fondamentaux de la recherche d'information**

### **La requête :**

La requête est l'interprétation d'un besoin en information de l'utilisateur exprimé à la machine et qui sera traitée par le système de recherche d'information ; c'est un ensemble de mots-clés. Elle peut être exprimée en langage naturel, booléen ou graphique.

### **Le document :**

Le document est l'élément central du système de recherche d'information, c'est un objet complexe sans cesse en évolution car il est lié aux développements des technologies de la communication [5].

La recherche des informations pertinentes qui répondent aux besoins d'un utilisateur, consiste à mettre en correspondance les représentations des informations contenues dans un fonds documentaire avec celles des besoins de l'utilisateur. Pour réaliser d'une façon efficace cette fonction, un système de recherche d'information doit réaliser :

- l'indexation des documents de la collection et du besoin utilisateur
- l'appariement des représentations requête documents pour le calcul de la pertinence des documents en réponse à un besoin utilisateur
- éventuellement, la reformulation de requêtes.

Pour résumer ces fonctions, nous pouvons représenter schématiquement certaines fonctionnalités d'un système de recherche d'information par ce que l'on appelle communément le processus « en U », tel que dans la figure I.1.

Ce processus manipule des collections de documents et des besoins utilisateur pour rechercher, à partir d'une collection de documents, ceux qui répondent au mieux à un besoin utilisateur. [5].

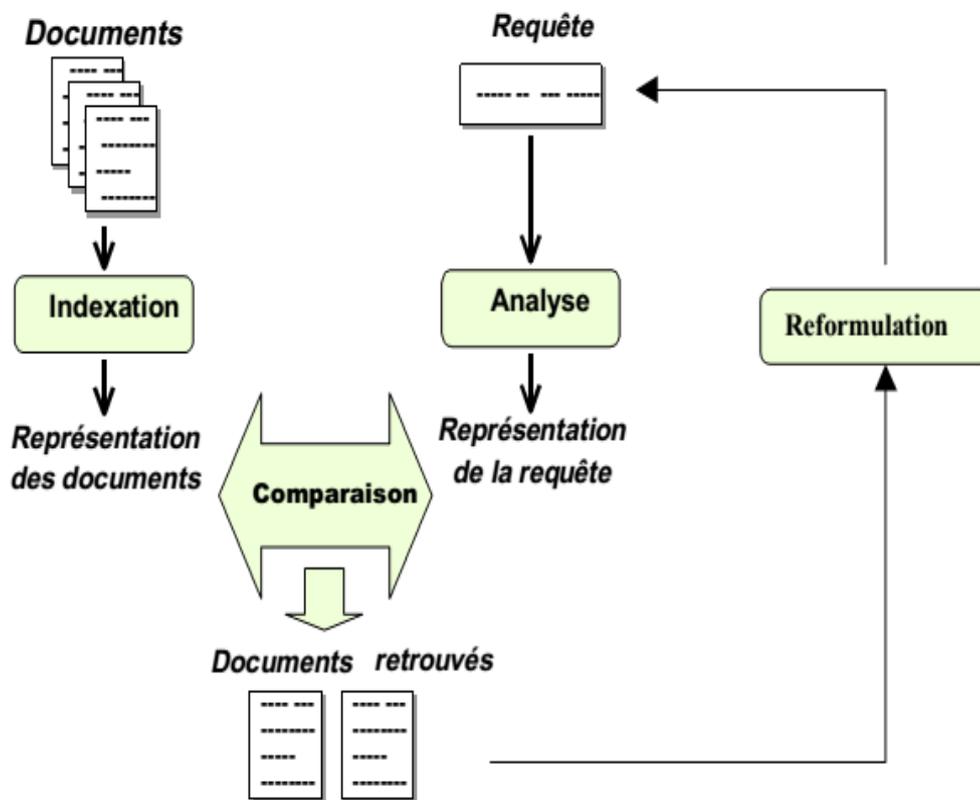


Figure I.1 : Processus en U de la RI [5]

## **I.6. Les types de systèmes de recherche d'information**

Les efforts continus des chercheurs en RI ont permis jusqu'à présent d'améliorer sans cesse les performances et la qualité des services d'accès à l'information. Dans ce paragraphe nous faisons un survol des principales évolutions dans le domaine de la RI de la première génération de systèmes de recherche d'information (SRI) dits classiques à la RI adaptative, puis récemment à la RI personnalisée [6].

### **I.6.1. Le système de recherche d'information classique**

Apparue dans les années soixante, a une vision orientée système, en ce sens où la recherche des informations pertinentes se base uniquement sur l'appariement des documents avec la requête soumise par l'utilisateur. Toutefois, cette vision de l'accès à l'information suppose que l'utilisateur est extérieur au système de recherche.

La différence de vocabulaire entre les termes choisis par l'utilisateur pour formuler sa requête et les termes utilisés pour représenter les documents engendrent un défaut d'appariement, Ce dernier est à l'origine d'une dégradation des performances de recherche et particulièrement dans le cas d'accroissement continu des sources d'information hétérogènes et la diversité des utilisateurs [6].

### **I.6.2. Le système de recherche d'information adaptatifs**

Les approches adaptatives des systèmes de recherche d'information exploitent les diverses sources d'évidence (documents jugés, termes pertinents, etc.) Pour aider et assister l'utilisateur à retrouver les informations pertinentes à son besoin.

### **I.6.3. Le Système de Recherche d'Information personnalisé**

Les approches dites de personnalisation exploitent les caractéristiques informationnelles spécifiques de l'utilisateur dans les processus d'accès à l'information [6].

## I.9. Défis d'un système de recherche d'information

On assiste aujourd'hui à une explosion de la quantité d'information disponible et accessible. En effet, l'évolution de la quantité d'information est exponentielle, d'après les dernières estimations.

On parle de plus de 180 millions de serveurs hôtes sur Internet et on a sûrement dépassé 2 milliards de pages en 2000. Si on regarde les rapports, Comme le montre la Figure I.2 [7].

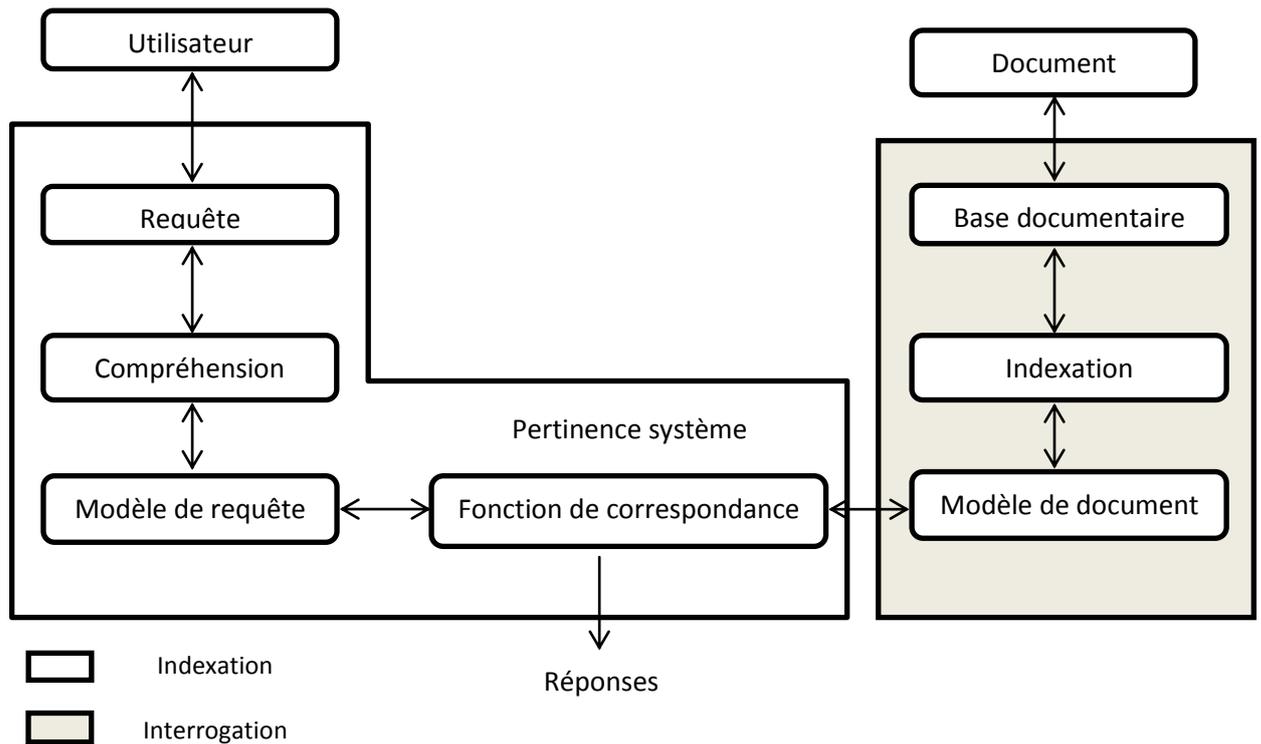


Figure I.2 : Système de recherche d'information [7]

## **I.10. Conclusion**

Dans ce chapitre, nous avons présenté les systèmes de recherche d'information, d'où nous avons cité les caractéristiques du besoin d'information, présenté les concepts fondamentaux de la recherche d'information.

Enfin, nous avons décortiqué les types des systèmes de recherche d'information (SRI classique, SRI adaptatifs, SRIP) et comme notre objectif est prise en compte de l'utilisateur dans la RI, nous allons présenter dans le chapitre suivant le profil utilisateur car le but d'un tel système est de satisfaire les besoin de utilisateur en besoin d'information.

## Chapitre II

### Le profil utilisateur

## **II.1. Introduction**

Le profil utilisateur y compris les informations qu'ils composent est un élément primordial dans le système de recherche d'information.

Plusieurs recherches ont montré que l'intégration de ce dernier dans le processus de recherche d'information ajoute beaucoup de bon résultat.

A cet effet, nous allons détailler cet aspect pour que nous pouvons l'introduire dans notre approche de recherche.

## **II.2. Notion de profil**

### **II.2.1. Le profil utilisateur**

Le profil utilisateur est une mise en facteur de la partie invariante des préférences de l'utilisateur. Même si l'ensemble des approches de personnalisation s'accordent sur l'importance de disposer d'un profil utilisateur.

Il manque un consensus sur la typologie des connaissances le constituant ainsi que sur la manière de représenter ces connaissances. Pour la majorité des approches, le profil contient la description du centre d'intérêt de l'utilisateur [8].

### **II.2.2. Le profil document**

Le document, élément central d'un système de recherche d'information, est un objet complexe sans cesse en évolution car il est lié aux développements des technologies de la communication.

La finalité d'un document est de transmettre de l'information. Mais cette information est un concept abstrait, porteur de sens. Ainsi l'information d'un même document, varie d'un individu à un autre selon la perception individuelle de chacun [8].

En outre, le sens d'un document est donné, non seulement par son contenu mais aussi par sa structure.

L'idée majoritaire aujourd'hui, est de renforcer la représentation de l'information du document en utilisant l'information sur l'information du document: les métas informations. On parle aujourd'hui de métadonnées et d'indexation « non thématique » en opposition à l'indexation « thématique » ou « conceptuelle » décrivant la sémantique de l'information [9].

L'objectif principal des métadonnées est de décrire, d'identifier et de définir une ressource. Toutefois, c'est dans leur rôle que repose le véritable intérêt des métadonnées: faciliter la RI.

Les parties de document ont un usage différencié à priori suivant le besoin de l'utilisateur. Dans la majorité des cas, un document a une structure générale qui forme une unité, car il est construit pour faire passer un message.

Cette unité matérielle et intellectuelle est le résultat d'un lien parfaitement établi entre ses différentes parties, celles-ci pouvant former à leur tour des unités indépendantes remplissant une fonction bien déterminée [9].

## II.3. Processus de définition du profil de l'utilisateur

Le profil de l'utilisateur couvre des aspects larges tels que son environnement cognitif, social et professionnel qui déterminent ses intentions au cours d'une session de recherche. La plupart des travaux actuels en RI contextuelle focalisent à juste titre, sur la représentation de l'aspect lié à ces intentions qualifiées de centres d'intérêts. Dans cette perspective, la modélisation du profil de l'utilisateur a pour objectif fondamental de représenter puis faire évoluer ses besoins en information à court et moyen terme. C'est une question qui pose la double difficulté de traduire les centres d'intérêt de l'utilisateur d'une part et faire émerger leur diversité d'autre part [9].

Le processus de définition du profil de l'utilisateur peut être caractérisé par trois phases. La première porte sur la représentation des unités d'information représentant le profil. La deuxième phase est liée à l'instanciation de ce modèle au cours d'une activité de recherche d'information.

Enfin, la troisième phase concerne l'évolution du profil au cours du temps. Chacune de ces phases met en jeu des approches et techniques de représentation et/ou de construction [9].

### II.3.1. Représentation

Le modèle de base le plus communément utilisé pour la représentation des centres d'intérêt de l'utilisateur est le modèle vectoriel où chaque centre est représenté par une liste de termes représentatifs. Cependant, on distingue trois principales approches de représentation : ensembliste, sémantique et multidimensionnelle [10].

- a) **Représentation ensembliste** : le profil y est généralement formalisé comme des vecteurs de termes pondérés ou classes de vecteurs non hiérarchisées ou hiérarchisées permettant de prendre en compte des centres d'intérêt multiples. La représentation non hiérarchique considère les centres d'intérêt comme indépendants dans la description du profil ; leur éventuelle dépendance peut être prise en compte lors de l'intégration du profil dans la phase d'évaluation de pertinence des documents. En revanche, la représentation sous forme de hiérarchie de classes permet de traduire les relations de spécificité/généralisation entre les centres d'intérêt.
- b) **Représentation sémantique** : la représentation du profil met en évidence, dans ce cas, les relations sémantiques entre informations le contenant. La représentation est essentiellement basée sur l'utilisation d'ontologies ou des réseaux sémantiques probabilistes.

Dans le cadre de cette approche, les centres d'intérêts de l'utilisateur sont appariés aux concepts des domaines de l'ontologie. Un profil est alors représenté en termes de concepts de l'ontologie intéressant l'utilisateur. Les ontologies de référence utilisées dans ce cadre sont basées sur la catégorisation en hiérarchie générale de Yahoo, Magellan, Lycos et ODP (Open Directory Project) [10].

- c) **Représentation multidimensionnelle:** le profil est structuré selon un ensemble de dimensions, représentées selon divers formalismes. Les propositions de standards P3P pour la sécurisation des profils ont défini des classes distinguant les attributs démographiques des utilisateurs (identité, données personnelles), les attributs professionnels (employeur, adresse, type) et les attributs de comportement (trace de navigation). Dans ce sens, on retrouve dans un modèle de représentation du profil structuré en dimensions (ou catégories) prédéfinies: catégorie de données personnelles, catégorie de données de la source, catégorie de données de livraison, catégorie de données de comportement et catégorie de données de sécurité. L'auteur a proposé ce modèle dans le cadre du développement d'un service avancé de bibliothèque numérique (recherche et livraison personnalisée de l'information sur le Web) à l'aide du système EUROgatherer [10].

### II.3.2. Construction

La construction du profil traduit un processus qui permet d'instancier sa représentation à partir de diverses sources d'information. Ce processus est généralement implicite et basé sur un procédé d'inférence du contexte et des préférences de l'utilisateur via son comportement lors de l'utilisation :

- d'un système d'accès à l'information : requêtes et documents jugés explicitement ou implicitement pertinents (consultés et/ou imprimés et/ou sauvegardés etc.)
- d'un navigateur web : liens explorés, dernières pages visitées etc.
- d'autres applications : les applications de bureautique, les outils de messagerie électronique etc. [10].

Les informations extraites de ces sources sont organisées selon le modèle de représentation du profil à l'aide de différentes techniques. La plus répandue est celle basée sur l'analyse statistique du texte selon l'algorithme de Rocchio.

L'autre technique largement utilisée également dans les procédés de construction du profil, est celle de la classification appliquée aux informations collectées de l'utilisateur. On distingue plusieurs variantes dont la classification simple, la classification hiérarchique ou la classification Bayésienne [10].

### II.4. Le contenu du profil

Chaque système de personnalisation propose un modèle de profil adapté au service qu'il offre. De façon formelle les données stockées dans les profils.

Dans cette section, nous allons voir la classification, l'organisation et la structuration des données des profils est un élément clé si on veut avoir une vision globale de la personnalisation. Différents travaux ont abordé cet aspect sans le couvrir dans son ensemble. Parmi ceux-ci le profil pour les utilisateurs d'une bibliothèque digitale composé de cinq catégories :

### II.4.1. Domaine d'intérêt

Le domaine d'intérêt est la dimension centrale du profil utilisateur. Cette dimension regroupe tous les attributs qui concernent les objets de contenu (informations ciblées). Elle peut définir aussi bien le domaine d'expertise et le niveau de qualification de l'utilisateur dans un domaine particulier que le contenu auquel s'intéresse l'utilisateur [10]. Le domaine d'intérêt peut être représenté de différentes manières Figure II.1:

- Par un vecteur de mots clés
- Par un ensemble de formules logiques
- Par un historique des interactions entre l'utilisateur et le système
- Par ontologies

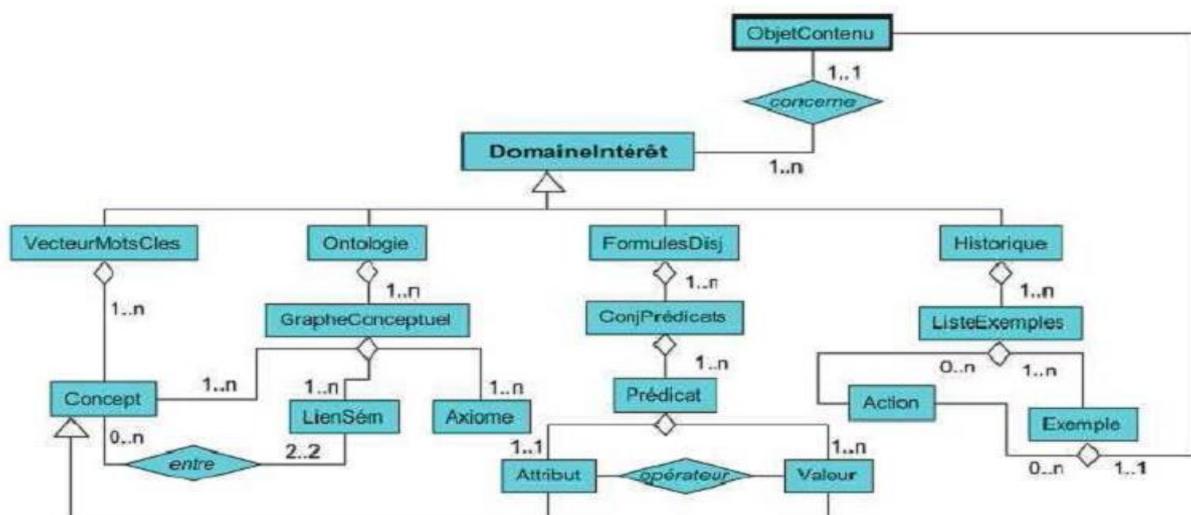


Figure II.1: Méta modèle de la dimension domaine d'intérêt [10]

### II.4.2. Données personnelles

Les données personnelles sont la partie statique du profil. Elles contiennent des informations qui décrivent l'utilisateur et ne dépendent pas du système à interroger. Ces informations sont classées selon trois catégories (Figure II.2).

1. Données personnelles : catégorie (identité) est composée d'un ensemble d'attributs d'identification de l'utilisateur. Des exemples de tels attributs sont le nom et le prénom de l'utilisateur, son adresse, son numéro de téléphone ou de fax, son adresse email etc.

2. Données démographiques : attributs démographiques comme par exemple la date de naissance de l'utilisateur, son genre, son revenu, son état civil, le nombre d'enfants etc.

3. Données des contacts : les contacts de l'utilisateur représentent son carnet d'adresse. Un élément de ces contacts est représenté par les données personnelles d'un autre utilisateur [10].

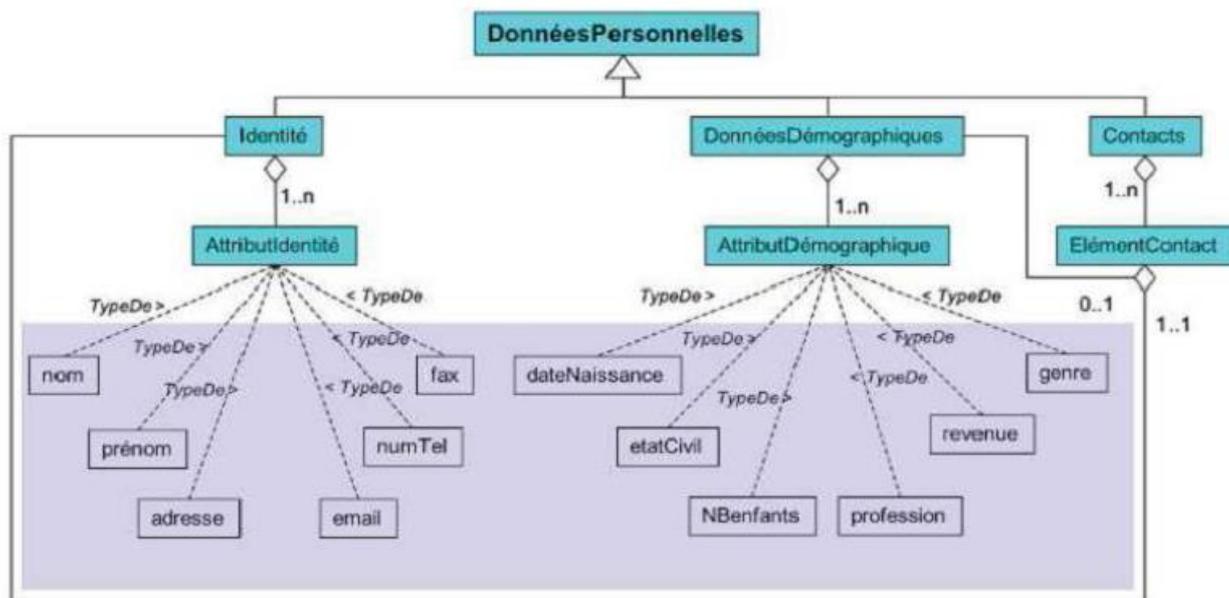


Figure II.2 : Méta modèle de la dimension donnée personnelles [10]

### II.4.3. Qualité

La dimension « Qualité » joue un rôle très important dans le domaine de la personnalisation . Les données de cette dimension décrivent la qualité attendue ou espérée par l'utilisateur qui sera confrontée à la qualité effective produite par le système de recherche d'informations afin de restreindre l'espace de recherche [10].

Les informations dans cette dimension sont représentées par des facteurs de qualité (Figure II.3).

Trois catégories de facteurs de qualité selon le type d'objets auxquels ils font référence sont distinguées :

1. Facteurs sur le contenu des données : cette catégorie concerne la qualité désirée des objets de contenu. Des exemples de facteurs de cette catégorie sont la fraîcheur et l'exactitude des données.
2. Facteurs sur le contenant des données : catégorie regroupe les facteurs de qualité des conteneurs des données (sources). La plupart des facteurs de cette catégorie résultent de l'expérience des autres utilisateurs (ex. le niveau de confiance dans la source, la fiabilité etc.). Ces facteurs sont utilisés pour filtrer les sources de données lorsque leur nombre est trop important.

- Facteurs sur les processus : délivrent ou notifient l'arrivée des résultats. Les facteurs de cette catégorie mesurent aussi bien les performances des processus (ex. temps de réponse) que leur capacité de produire tous les résultats pertinents (rappel) et uniquement les résultats pertinents (précision) [10].

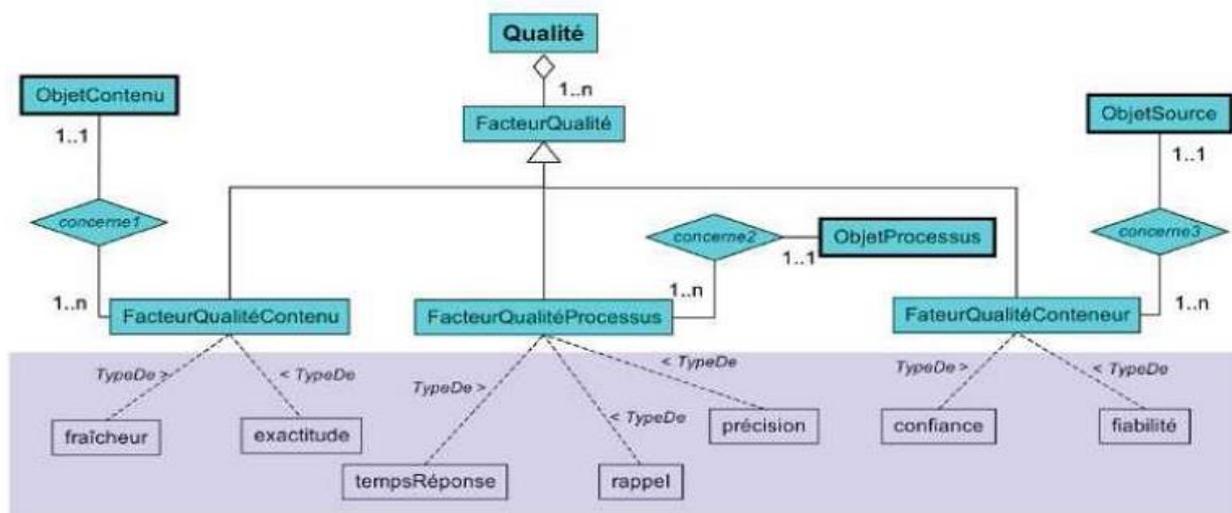


Figure II.3: Méta modèle de la dimension Qualité [10]

#### II.4.4. Données de livraison

Les données de livraison concernent d'abord tout ce qui est lié aux modalités de présentation des résultats en fonction de la plateforme de l'utilisateur, de la nature et du volume des informations délivrées, des préférences esthétiques ou visuelles de l'utilisateur. À ces modalités de présentation, on peut ajouter les modalités de livraison et de notification, décrivant le moment et le moyen de livraison des résultats et la manière de notifier cette livraison (différé, immédiat par exemple) [10].

Les données de livraison sont divisées en deux classes (Figure II.4) :

- des caractéristiques des objets de contenu délivrés: le format des documents, leur nombre et leur taille ou encore la mise en page des données ;
- des informations sur les interactions entre le système et l'utilisateur : des informations sur le moment et le moyen des interactions entre l'utilisateur et le système [10].

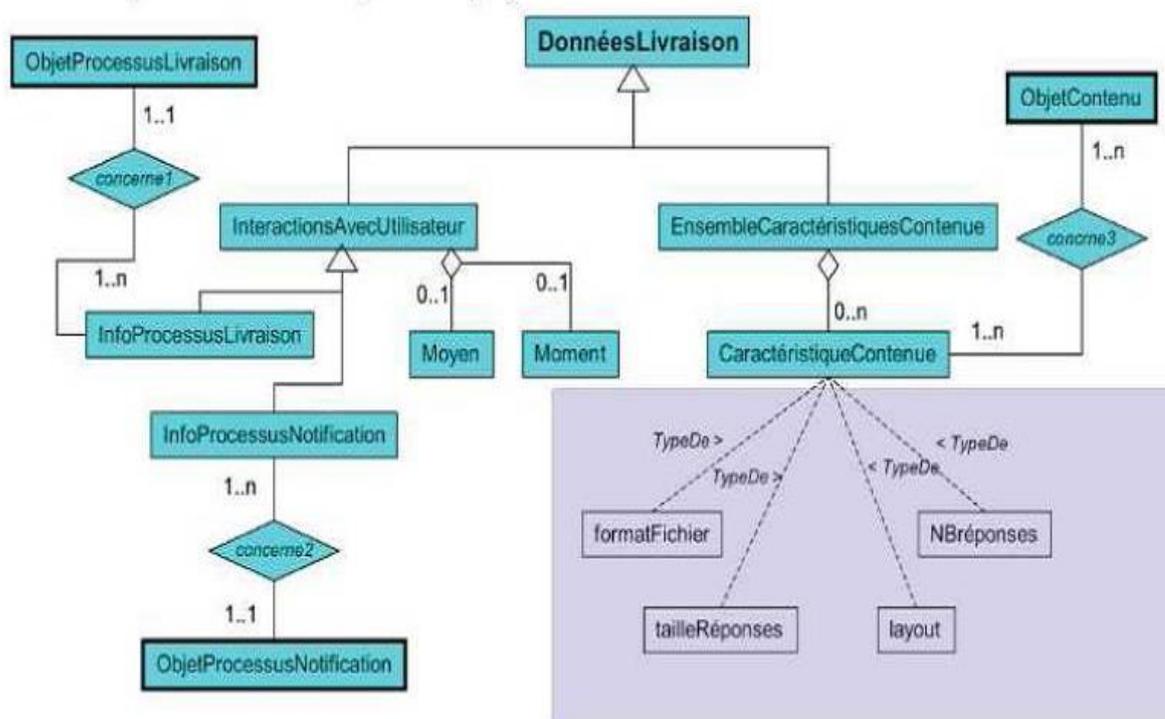


Figure II.4: Méta modèle de la dimension donnée de livraison [10]

### II.4.5. Données de sécurité

La sécurité est une dimension fondamentale du profil. Elle peut concerner les données que l'on interroge ou modifie, les informations que l'on calcule, les requêtes utilisateurs ou les autres dimensions du profil [10].

Trois types de sécurité en fonction des objets concernés sont distingués (Figure II.5) :

1. la sécurité du profil utilisateur : se fait à travers des autorisations d'accès aux informations du profil. Les utilisateurs d'un profil, qui peuvent être le système d'information ou les autres utilisateurs particuliers, se voient attribuer des droits d'accès ou des paquets de droits d'accès (rôle) afin de pouvoir interroger les données de ce profil.
2. la sécurité des résultats : concerne les objets de contenu qui sont délivrés à l'utilisateur. Elle est exprimée en indiquant le niveau de sécurisation des résultats (sur une échelle prédéfinie) ou à travers des expressions dans un formalisme existant .
3. la sécurité du processus de production des résultats : exprime la volonté de l'utilisateur de cacher un traitement sensible qu'il effectue [10].

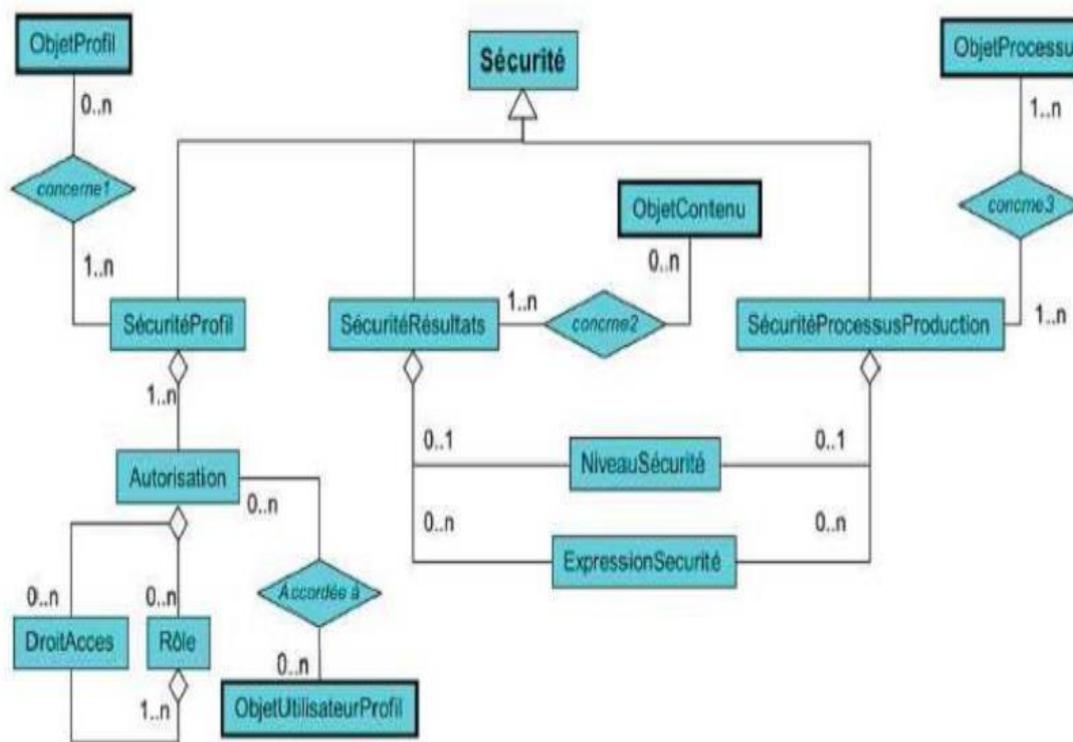


Figure II.5 : Méta modèle de la dimension sécurité [10].

## II.5. Exemple d'un système de recherche d'information utilisant le prof : SRIP

Un système de recherche d'information personnalisé (*SRIP*) est un système qui intègre l'utilisateur, en tant que structure informationnelle, tout au long de la chaîne d'accès à l'information [11].

Le *SRIP* ne se limite pas seulement à modéliser les caractéristiques des utilisateurs en des profils. Il doit être capable de déduire à partir de ces profils, l'intention de l'utilisateur lorsqu'il effectue sa recherche, en d'autres termes son contexte de recherche, et de détecter l'évolution des profils de manière dynamique. Le système doit donc inclure :

- des techniques et algorithmes pour capturer et modéliser le but, les préférences et les centres d'intérêts de l'utilisateur ou un groupe d'utilisateurs. Un modèle de profil utilisateur est alors décrit et instancié,
- une procédure de mise à jour du profil qui traduit son évolution dans le temps,
- des mécanismes et algorithmes pour intégrer le profil de l'utilisateur dans le processus d'accès et retourner l'information pertinente en fonction de ce profil.

[11] Comme le montre la Figure II.6.

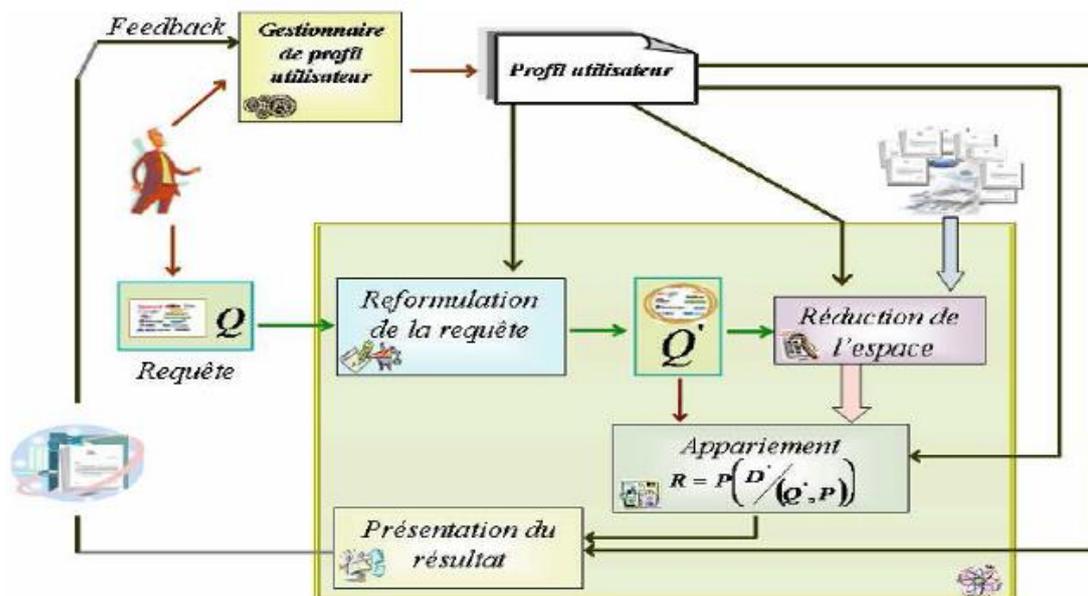


Figure II.6: Processus global d'accès personnalisé à l'information [11].

La diversité des systèmes de personnalisation rend la définition d'une architecture fonctionnelle et formelle d'accès personnalisé à l'information difficilement généralisable. Néanmoins, nous tentons dans ce paragraphe de dégager une architecture standard pour un système d'accès personnalisé à l'information, présentée par la Figure précédant.

Notre volonté est de mettre en évidence l'ensemble des fonctionnalités de tels systèmes même si elles ne sont pas toujours présentes collectivement dans tout système [11].

▪ Cette architecture centrée autour de l'utilisateur met en évidence :

1. Un gestionnaire de profil pour représenter, construire et faire évoluer les profils des utilisateurs.

2. Les étapes du cycle de vie de la requête où l'on intègre le profil utilisateur dans :

(a) la phase de reformulation de la requête afin de mieux cibler le contexte de la recherche de l'utilisateur.

(b) la phase de réduction de l'espace de recherche pour restreindre l'espace de recherche aux documents qui ciblent les besoins de l'utilisateur.

(c) la phase d'appariement pour calculer la pertinence des documents en fonction des caractéristiques spécifiques de l'utilisateur.

(d) la phase de présentation des résultats pour restituer les informations selon le contexte et les préférences de l'utilisateur [11].

### II.5.1. Modélisation du profil utilisateur

L'introduction de la dimension utilisateur dans un processus d'accès à l'information, mérite, voire nécessite une réflexion sur la modélisation de l'entité utilisateur.

La fiabilité ou qualité des profils est en effet d'une importance bien connue dans le domaine de la modélisation utilisateur. En effet, on constate que l'une des principales raisons du manque de performances des techniques de personnalisation est typiquement l'application d'un profil utilisateur hors contexte [12].

Les utilisateurs peuvent avoir des préférences générales, récurrentes et stables. Cependant, l'ensemble des informations contenues dans le profil ne sont pas forcément appropriées à toutes les situations de recherche. Le plus souvent, les systèmes n'utilisent seulement qu'un sous-ensemble de ces informations, qu'ils supposent pertinents pour la recherche en cours.

Dès lors, le choix du profil adéquat constitue la principale réflexion lors de la mise en œuvre du SRIP. De ce fait, les questions fondamentales posées pour modéliser le profil utilisateur sont le :

- Quoi?
  - Quelles propriétés informationnelles caractérisent l'utilisateur ?
  - Quelle structure informationnelle utiliser pour représenter l'utilisateur ?
- Comment?
  - Comment collecter les informations du profil ?
  - Comment construire le profil de l'utilisateur ?
  - Comment détecter le contexte, le but de la recherche et les besoins à court/long
  - termes de l'utilisateur ?
  - Comment adapter le profil à l'évolution de l'utilisateur lui-même ?
  - Comment assurer la sécurité et la confidentialité des informations du profil ?
- Quand ?
  - Quand faut-il faire évoluer le profil de l'utilisateur ? [12].

## II.5.2. Exploitation du profil d'utilisateur

Le domaine d'application conditionne fortement les techniques de personnalisation employées par le système. Cela a en effet un impact direct sur l'exploitation du profil dans la chaîne d'accès à l'information et par conséquent sur les mécanismes mis en œuvre. Les principales interrogations posées concerne le « Quoi » le « Comment » de la mise en œuvre :

- Quoi ?
  - Quels services de personnalisation proposée : de la recommandation et/ou du filtrage, de l'aide à la navigation, un assistant personnel de recherche ?
  - Dans quelles étapes du cycle de vie de la requête faut-il intégrer le profil ?
  - Quelles informations du profil exploiter lors de l'accès à l'information ?
- Comment ?
  - Comment intégrer le profil de l'utilisateur dans le processus de personnalisation ?
  - Comment évaluer l'impact de la personnalisation sur le processus de recherche ? [13].

## II.6. Conclusion

Nous étions dans cet axe aux notions de profil et le contenu de celui-ci, qui varient en fonction de l'objectif et le sujet de qui a été fait pour lui, ou l'importance de la relation avec le système de l'utilisateur, la recherche d'information et les processus définis profil de l'utilisateur, et les problèmes les plus importants auxquels nous sommes confrontés dans la modélisation et l'exécution profil de l'utilisateur.

Dans le prochain chapitre, nous allons présenter la modélisation du profil utilisateur.

## Chapitre III

### La modélisation du profil utilisateur

### III.1 Introduction

La représentation de l'utilisateur à travers la notion de profil permet de mieux comprendre certains mécanismes cognitifs, notamment ceux permettant de percevoir le concept subjectif de la pertinence et au-delà, cibler ses besoins spécifiques dans le but d'améliorer les performances de recherche. Un modèle de représentation permet d'organiser ces éléments afin de faciliter leur exploitation dans le processus d'accès à l'information.

Un des problèmes les plus difficiles à résoudre dans la personnalisation consiste à trouver des techniques de construction et de mise à jour des paramètres du profil d'un utilisateur.

La façon la plus naturelle est de remplir manuellement les cases des paramètres du profil. Bien que l'utilisateur soit celui qui connaît le mieux ses préférences, il est parfois incapable de les exprimer et les définir de façon formelle. De plus, le processus de saisie de tous les paramètres du profil est souvent long et ennuyeux.

Afin de résoudre ces problèmes, ils existent plusieurs méthodes de capture des paramètres du profil de façon automatique. Ces méthodes sont liées l'interprétation des activités de l'utilisateur sur les données du résultat.

Le principe de base de ces techniques est l'étude du comportement de l'utilisateur et la classification de ses caractéristiques ou celles des objets recherchés, Grâce au suivi de l'utilisateur, elles permettent la mise à jour de son profil.

## III.2 Le filtrage d'information collaboratif pour de modélisation du profil

### III.2.1 Filtrage de l'information

Le filtrage de l'information pourrait être défini comme étant le processus qui vise à filtrer les informations d'un flux pour ne faire parvenir que celles qui intéressent l'utilisateur, contrairement à la recherche d'information où l'utilisateur à l'aide d'une requête sélectionne l'information pertinente à partir du flux. Les propriétés principales d'un système de filtrages sont :

- Filtrer l'information non désirée pour ne garder que ce qui est intéressant.
- Le flux d'information est dynamique (envoi d'alertes, ...).
- Personnalisation via la constitution et la maintenance d'un profil utilisateur, et prise en charge des intérêts à long terme.
- L'adéquation entre besoin et ressource peut être fait avec plusieurs dispositifs : indexation, contenu sémantique, évaluation, ...
- Interaction de l'utilisateur, avec l'évaluation des ressources recommandées, pour la mise à jour de son profil [13].

Plusieurs systèmes de filtrage existent, initialement appliqués aux forums, e-mails, ... leur domaine d'application s'est étendu pour inclure les ressources documentaires en général sans oublier que ces systèmes sont largement utilisés dans les domaines du commerce électronique (Amazon) et des divertissements (MyStrands, Last.fm, Netflix, Movielens, ...).

Le filtrage d'information est catégorisé en trois familles :

- **Filtrage cognitif** : basé sur le contenu.

Le filtrage basé sur le contenu (figure III.1) peut être vu comme un système de recherche d'information où la requête joue le rôle d'un filtre permanent ou à long terme. Ce type de systèmes est basé sur un processus à deux phases :

- La première étant la sélection de ressources correspondant au profil de l'utilisateur,
- La deuxième étant la mise à jour du profil de l'utilisateur après retour de pertinence des résultats [13].

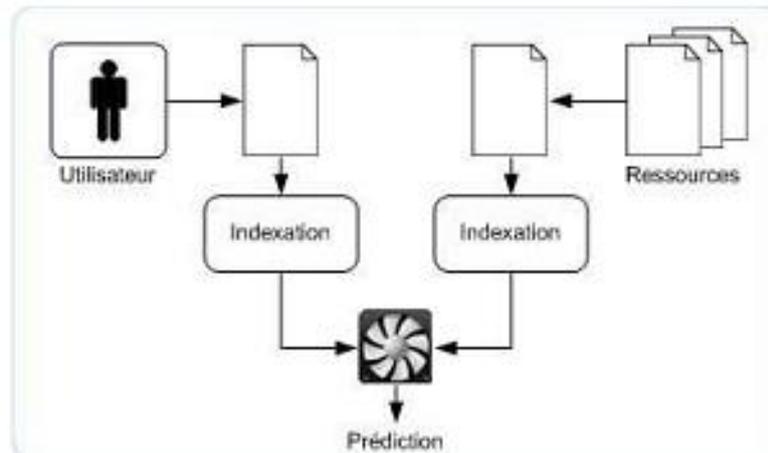


Figure III.1: Filtrage basé sur le contenu [13]

- **Filtrage collaboratif** : basé sur les évaluations des utilisateurs sur les ressources.

le filtrage collaboratif (figure III.2 ) se base sur le principe qui dit qu'on devrait pouvoir se servir de ce que les autres ont déjà trouvé et évalué pour trouver ce que nous cherchons. C'est comme la notion de bouche à oreille sur une grande échelle. Le fait que ce type de systèmes ne soit pas basé sur le contenu, permet de filtrer n'importe quel type de ressources qu'elles soient : textuelle, image, vidéo et même physique. Elle est au contraire basée sur la construction de profils d'utilisation, ces profils seront créés au fur et à mesure par les évaluations des utilisateurs [13].

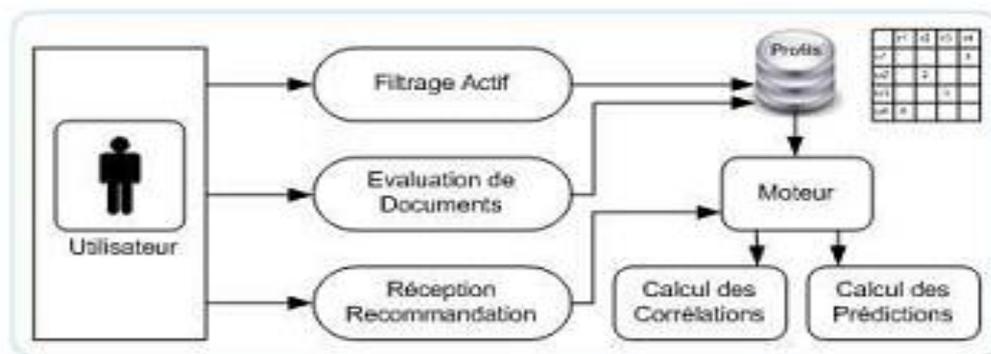


Figure III.2: Processus de filtrage collaboratif [13]

- **Filtrage hybride** : Combine les deux approches.

Constatant les avantages et inconvénients des deux approches, on comprend naturellement l'utilité des méthodes hybrides qui visent à combiner les deux types d'approches pour limiter les inconvénients. L'hybridation des méthodes peut être effectuée avec différentes méthodes, parmi lesquelles : la pondération, le changement adaptatif, la composition, la combinaison, la cascade, l'augmentation de fonctionnalité, etc...[13].

### III.2.2 Modélisation d'un profil

La modélisation d'un profil utilisateur (figure III.3) se traduit en l'énumération des informations nécessaires à la description de cet utilisateur. L'objectif de la modélisation est la mise en place d'une structure qui contiendra les informations liées aux caractéristiques de l'utilisateur .

Différents travaux ont abordé la structuration et la modélisation du profil utilisateur, le W3C a défini le P3P pour la sécurisation des profils, 3 classes sont définies : attributs démographiques (identité et infos personnelles), attributs professionnels (employeur, adresse, type) et des attributs de comportement (trace de navigation) [13].

Le chercheur Amato en 1999 propose 5 catégories pour les utilisateurs d'une bibliothèque digitale : données personnelles (identité), données collectées (contenu, structure et provenance des documents), données de livraison (temps et moyen de livraison), données de comportement (interaction de l'utilisateur avec le système), données de sécurité (conditions d'accès).

Les travaux des chercheurs Kostadinov en 2004 et Bouzeghouben 2004 considèrent que les approches citées précédemment sont insuffisantes pour couvrir le contexte de la personnalisation dans les systèmes de filtrage et de recommandation [13]. Ils distinguent 8 dimensions, 7 auxquelles on peut ajouter une huitième contenant toutes les informations inclassables.

- Données personnelles
- Centres d'intérêt
- Ontologie du domaine
- Qualité attendue des résultats délivrés
- Customisation
- Sécurité et confidentialité
- Retour de préférences (feedback)
- Informations diverses

La dimension ‘Données personnelles’ est caractérisée par : l’identité de la personne, l’âge, le sexe, la profession, les langues parlées, ...

Les centres d’intérêt sont en général décrits sous la forme de mots clés. L’ontologie du domaine permet de rendre explicite la sémantique de certains termes employés comme centres d’intérêt.

La qualité attendue exprime le niveau de qualité, l’origine, la précision, la fraîcheur et la crédibilité de l’information [13].

La customisation concerne l’adaptation du résultat selon les commodités d’utilisation, l’esthétique, etc. On y trouve des données sur l’environnement de l’utilisateur comme :

- La localisation géographique de l’utilisateur
- Des paramètres temporels
- le matériel et logiciels utilisés pour accéder à l’information.
- Performance, nombre de résultats, etc.

Ces données ont la propriété qui leur permet d’être trouvées automatiquement, exemple le ‘User Agent’ qui permet de définir le logiciel/matériel utilisé, l’adresse IP permet de trouver le pays voire la localisation, un système de géo localisation peut déterminer votre emplacement exacte, etc.

La sécurité et confidentialité concerne la définition des droits d’accès au système et l’expression de la volonté de cacher certaines actions faites par l’utilisateur. Le retour de préférences désigne le feedback de l’utilisateur concernant les ressources, qu’il soit explicite ou bien implicite en analysant le comportement de l’utilisateur sur le système [13].

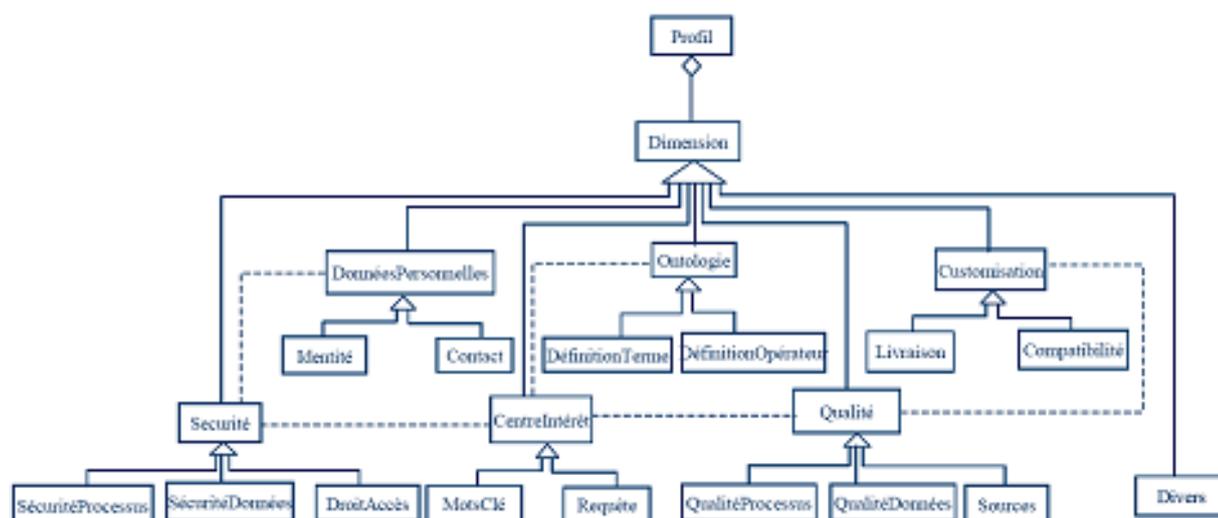


Figure III.3: Modélisation d'un profil utilisateur [13]

### III.3 Techniques statistiques pour la définition du profil d'un utilisateur

La mise en œuvre naïve d'un modèle classique de recherche d'information suppose que l'utilisateur est complètement représenté par sa requête et que les résultats retournés pour une même requête sont identiques même si elle est exprimée par des utilisateurs différents. Les problèmes immédiats posés par une telle hypothèse sont notamment l'ambiguïté du sens des mots, l'impossibilité de sélectionner des sources opportunes et l'inintelligibilité des résultats. En outre, ces problèmes sont d'autant plus accentués que les requêtes sont courtes ( $\approx 2.29$  mots par requête) et que les sources d'information sont volumineuses et hétérogènes. Les premières solutions apportées à ce type de problèmes et pouvant s'apparenter à la personnalisation, sont les techniques de reformulation de requêtes par injection de pertinence. Cependant, vu le contexte actuel lié au volume d'informations, ces techniques sont peu fiables. Les travaux s'orientent actuellement vers une définition plus large de l'utilisateur permettant de l'exploiter dans la chaîne d'accès à l'information. Cette section présente les notions ayant émergé de cette direction de travaux puis en rapporte les principales contributions portant sur la définition du profil de l'utilisateur [11].

Ces techniques sont basées essentiellement sur des modèles théoriques issus de la statistique soutenus par des heuristiques et des algorithmes appropriés. Citons parmi elles:

- **Modèle linéaire** : Le modèle linéaire a une structure simple. L'hypothèse de base est que la valeur de prédiction présumée et inconnue d'un objet cible du système par exemple, un article à recommander de chercheur Raskutti en 1997 est une combinaison linéaire des valeurs calculées à partir d'un comportement passé de l'utilisateur.  
Le modèle linéaire peut être aisément combiné avec des techniques collaboratives ou les valeurs connues sont issues de l'appréciation des membres des groupes associés à l'utilisateur courant.
- **Modèle Markovien** : Ce modèle est essentiellement basé sur l'hypothèse markovienne qui permet de représenter une séquence d'événements ultérieurs sur la base d'un nombre fixe d'événements antérieurs. Etant donnée la distribution de probabilités d'occurrences des événements passés, la théorie markovienne offre alors des éléments pour calculer la probabilité d'occurrence des événements futurs [11].

- **Réseaux de neurones** : les réseaux de neurones sont des structures basées sur l'interconnexion de nœuds et un principe d'activation par propagation de signaux à travers les connexions depuis les entrées jusqu'aux sorties. La signification effective des nœuds, des connexions et des valeurs d'activation dépend du problème pour lequel est dédié le réseau. De manière générale, les réseaux de neurones sont destinés à résoudre des problèmes de décision non linéaires. Dans le cas de la modélisation utilisateur, l'entrée représente une situation ou faits observables à partir de l'utilisateur, les sorties représentent des objets cible du système avec des valeurs d'activation qui traduisent le degré de prédiction [11].
  
- **Classification** : les méthodes de classification permettent de partitionner un espace d'objets en classes de manière à réduire sa dimension. Les objets d'une même classe ont des propriétés partageables dont le degré de corrélation est calculé à l'aide de métrique basées sur la similarité. De nombreuses stratégies de classification sont proposées dans la littérature. Du point de vue de la modélisation utilisateur, les méthodes de classification permettent généralement d'identifier la classe de caractéristiques de l'utilisateur courant à partir d'informations dérivées de son comportement [11].
  
- **Induction de règles** : l'induction de règles consiste en l'apprentissage de règles de prédiction à partir d'un ensemble de faits observés. Contrairement aux méthodes de classification, les techniques d'induction requièrent, durant l'apprentissage du système, l'identité de la classe associée à chaque observation. Ces techniques produisent un ensemble de règles proprement dites ou des arbres de décision.
  
- **Réseaux bayésiens** : les réseaux bayésiens sont des graphes acycliques orientés où les nœuds correspondent à des variables aléatoires. Les nœuds sont interconnectés à l'aide de liens orientés qui représente des liens de causalité entre nœuds parent et nœuds fils. A chaque nœud est associée une distribution de probabilité conditionnelle qui permet d'assigner au nœud une valeur de probabilité dépendante de la combinaison des valeurs de probabilités possibles des nœuds parents. Ces dernières années, les réseaux bayésiens ont largement investi les travaux sur la modélisation utilisateur. Ils sont plus flexibles que les techniques précédentes, puisqu'ils permettent de représenter explicitement les relations de causalité entre faits et d'émettre des prédictions sur de nombreux paramètres du système [11].

## III.4 Techniques base sur l'ontologie

### III.4.1 Notion d'ontologie

Le terme Ontologie est un terme grec composé des mots (figure III.4): « ontos » et « logos » qui veulent dire respectivement l'essence de l'être. À l'origine, l'Ontologie est une branche de la philosophie dans laquelle les philosophes ont tenté de rendre compte de l'existant de façon formelle. Ce terme, qui hérité d'une tradition philosophique, est apparu dans le domaine informatique grâce au projet ARPA Knowledge Sharing Effort en 1991 [11].

En informatique, "ontologie" est un objet et non une science. C'est le résultat d'une tentative de formulation exhaustive et rigoureuse de la conceptualisation d'un domaine.

Les ontologies sont apparues au début des années 90 dans la communauté Ingénierie des connaissances, dans le cadre des démarches d'acquisition des connaissances pour les systèmes à base de connaissances (SBC) . Les ontologies sont étudiées aussi par les chercheurs travaillant en intelligence artificielle, sur la représentation de la connaissance, et maintenant, sur le Web sémantique [11].

Les ontologies sont utilisées, entre autres, pour représenter les données disponibles sur le Web. Les motivations sont la réutilisation, le partage et la communication des connaissances entre systèmes, afin d'intégrer de multiples sources d'information. Un de leurs buts est de faciliter la recherche et l'extraction de connaissances en permettant d'inférer des connaissances.

La construction d'une ontologie permet de clarifier les concepts d'un domaine, son but est d'éviter autant que possible les ambiguïtés sémantiques F. Gandon, décrit ainsi ce domaine de recherche : « l'ontologie est un nouvel objet de l'intelligence artificielle, qui a récemment atteint sa maturité en devenant un puissant outil conceptuel pour la modélisation de connaissances » [11].

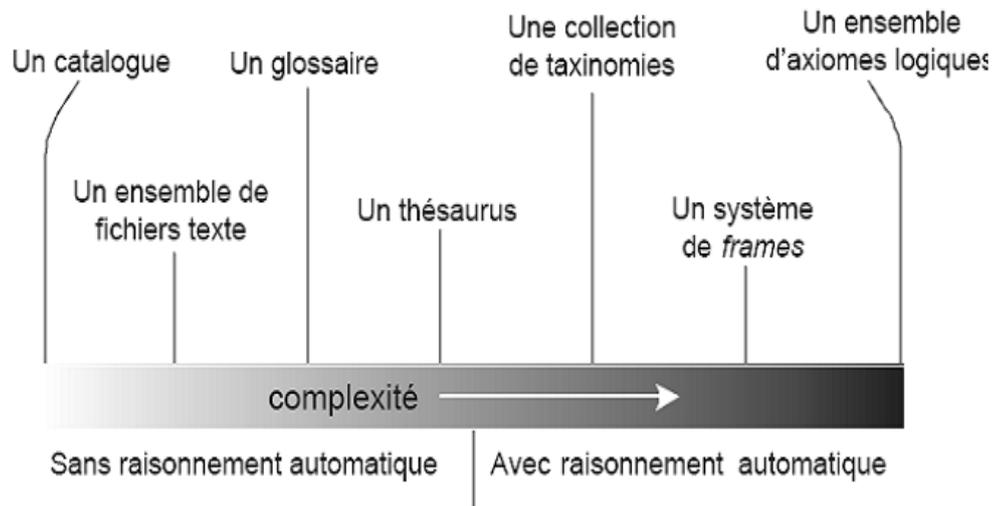


Figure III.4: Différentes acceptations du terme ontologie [11]

**Une ontologie:** vocabulaire, relations, propriétés, limitations des valeurs, spécifications des disjonctions et des relations logiques (figure III.5).

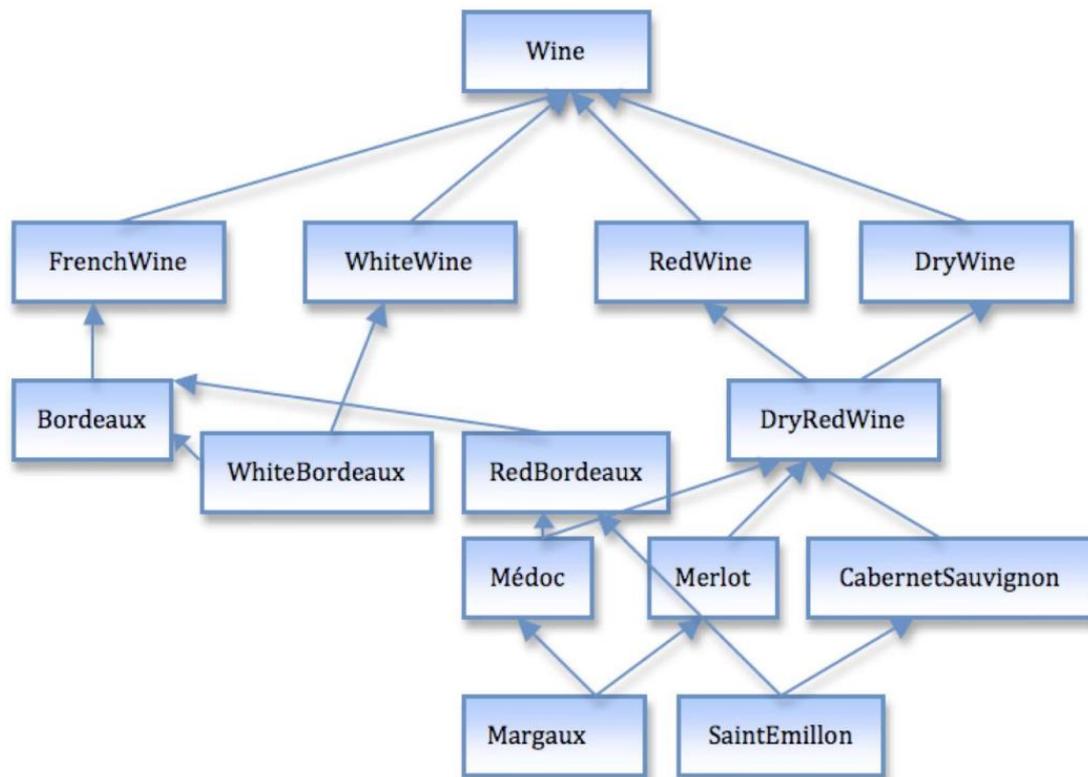


Figure III.5 : schéma de l'ontologie [11]

### III.4.2 Cycle de vie d'une ontologie

Étant donné que les ontologies seront utilisées comme composantes de systèmes logiciels, leur développement doit s'appuyer sur les mêmes principes que ceux utilisés en génie logiciel. En particulier, elles doivent être considérées comme des objets techniques évolutifs et posséder un cycle de vie spécifique. La construction d'une ontologie doit passer par différentes phases, à commencer par l'analyse des besoins, jusqu'à la validation et l'évaluation. [11], comme le montre la figure suivante:

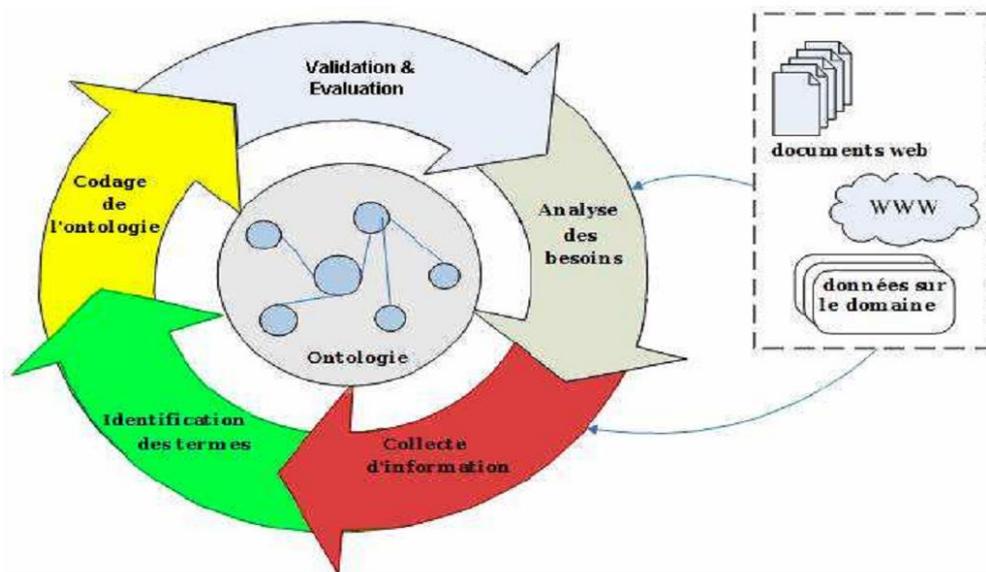


Figure III.6 cycle de vie d'une ontologie [11]

### III.4.3 Principes de développement d'une ontologie

La construction d'une ontologie suppose certaines obligations qui découlent du choix d'utiliser certains concepts plutôt que d'autres pour représenter un phénomène.

En effet, il existe un ensemble de critères et de principes qui ont fait leurs preuves dans le développement des ontologies et qui peuvent être résumés en ce qui suit,

- **Clarté** : les ambiguïtés doivent être réduites. Quand une définition peut être axiomatisée, elle doit l'être. Dans tous les cas, des définitions en langage naturel doivent être fournies.
- **Cohérence** : une ontologie doit être cohérente. Les axiomes doivent être consistants. La cohérence des définitions en langage naturel doit être vérifiée autant que faire se peut.
- **Extensibilité** : l'ontologie doit être construite de telle manière que l'on puisse l'étendre facilement, sans remettre en cause ce qui a déjà été fait.

- **Biais d'encodage minimal** : l'ontologie doit être conceptualisée indépendamment de tout langage d'implémentation. Le but est de permettre le partage des connaissances (de l'ontologie) entre différentes applications utilisant des langages de représentation différents.
- **Engagement ontologique minimal** : une ontologie doit faire un minimum d'hypothèses sur le monde : elle doit contenir un vocabulaire partagé mais ne doit pas être une base de connaissances comportant des connaissances supplémentaires sur le monde à modéliser [11].

### III.5 Les composants d'une ontologie

Une ontologie peut être vue comme un treillis de concepts et de relations entre ces concepts destinés à représenter les objets du monde sous une forme compréhensible aussi bien par les hommes que par les machines. Une ontologie est constituée de concepts et des relations ainsi que des propriétés et des axiomes [11].

- Les concepts sont des notions (ou objets) permettant la description d'une tâche, d'une fonction, d'une action, d'une stratégie ou d'un processus de raisonnement, etc. Ils peuvent être abstraits ou concrets, élémentaires ou composés, réels ou fictifs. Habituellement, les concepts sont organisés en taxinomie. Une taxinomie est une hiérarchie de concepts (ou d'objets) reliés entre eux en fonction de critères sémantiques particuliers.
- Les relations sont les liens organisant les concepts de façon à représenter un type d'interaction entre les concepts d'un domaine. Elles sont formellement définies comme un sous ensemble d'un produit de  $n$  ensembles. C'est à dire  $R: C_1 \times C_2 \times \dots \times C_n$ . Des exemples de relations binaires sont : sous-concept-de, connect-à, sorte de, etc... [11].
- Les propriétés sont des restrictions des concepts ou des relations.
- Les fonctions sont des cas particuliers de relations dans lesquelles le  $N$ ème éléments de la relation est unique pour les  $n-1$  précédents. Formellement, les fonctions sont définies ainsi  $F: C_1 \times C_2 \times \dots \times C_{n-1}, C_n$ .
- Les axiomes de l'ontologie permettent de définir la sémantique des termes (classes, relation), leurs propriétés et toutes contraintes quant à leur interprétation. Ils sont définis à l'aide de formules bien formées de la logique du premier ordre en utilisant les prédicats d'ontologie.
- Les instances sont utilisées pour représenter des éléments [11].

### III.6 Méthodologie de l'ingénierie ontologique

Les méthodologies de l'ingénierie ontologique présentées par Fernandez et Corcho, aident l'utilisateur à construire des ontologies mais elles restent insuffisantes et aucune d'elles n'est standardisée. Mais, elles se rejoignent sur un processus général de construction d'ontologie [11]. Cette dernière repose sur un enchaînement de quatre étapes : (figure III.7).

- **La conceptualisation**, sert à dégager à partir des données brutes les concepts et les relations entre ces concepts permettant de décrire de manière informelle les entités cognitives du domaine.
- **L'ontologisation**, est la formalisation partielle, indépendante des contraintes d'opérationnalisation, du modèle conceptuel en précisant la sémantique à l'aide d'axiomes et de contraintes sur les concepts et les relations. elle est semi-formelle, certaines connaissances ne pouvant pas forcément être formalisées à ce niveau.
- **L'opérationnalisation**, est une formalisation de l'ontologie conceptuelle à l'aide d'un langage de représentation de connaissances pour obtenir une représentation formelle des connaissances du domaine, basé sur l'ontologie.
- **L'implémentation**, est le passage du modèle formel de connaissances à son implémentation en machine. Cette étape peut imposer des contraintes lors de l'étape précédente [11].

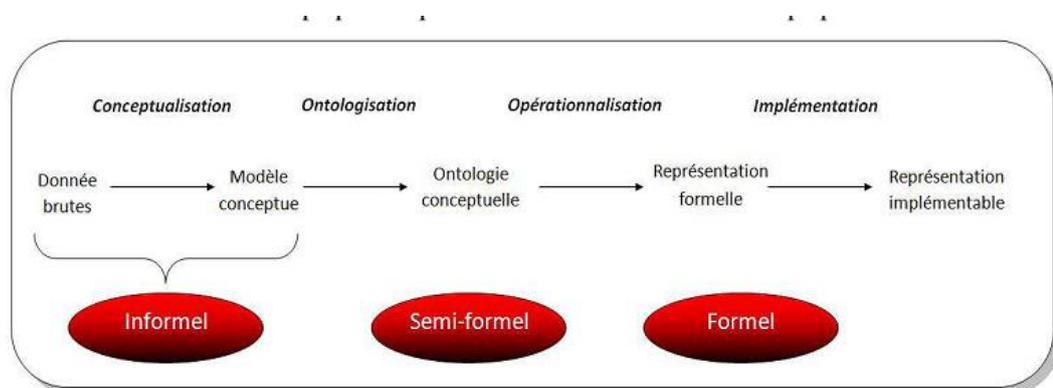


Figure III.7: Processus de construction d'ontologie [11].

### III.7 Les types d'ontologies

Les types d'ontologies les plus utilisés sont :

- **Les ontologies générales:** elles portent sur les concepts généraux qui se veulent indépendants d'un domaine ou d'un problème particulier, tels que les concepts de temps, d'espace, de notions mathématiques.
- **Les ontologies du domaine:** ce type exprime des conceptualisations spécifiques à des ontologies de domaines particuliers tout en étant générique pour ce domaine. Ces conceptualisations mettent des contraintes sur la structure et les contenus des connaissances du domaine.
- **Les ontologies d'application:** ces ontologies sont les plus spécifiques. Les concepts correspondent souvent aux rôles joués par les entités du domaine tout en exécutant une certaines activités. Elles peuvent contenir des extensions spécifiques telles les méthodes et les tâches. Elles contiennent toutes les définitions nécessaires pour décrire la connaissance requise pour une application particulière [12].

### III.8 Modélisation du profil utilisateur base sur l'ontologie

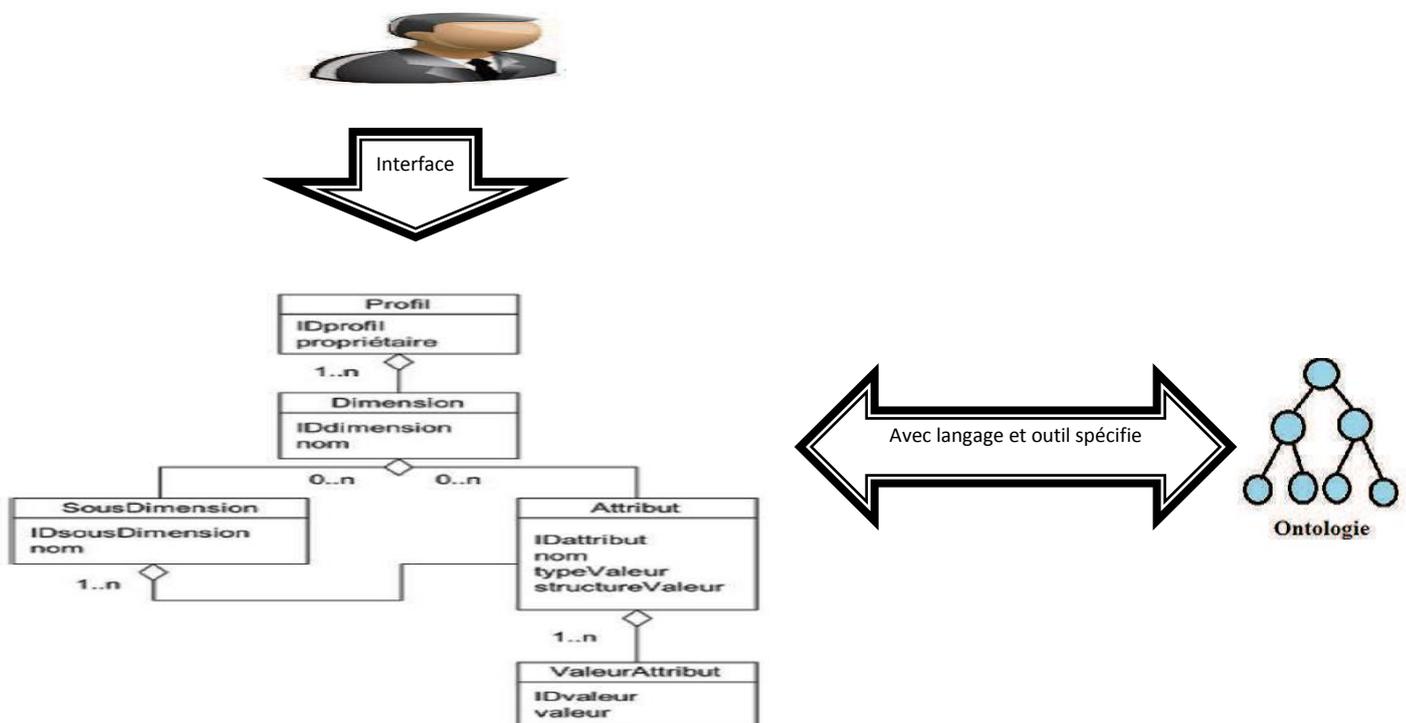


Figure III.8: Modélisation du profil utilisateur

Le modèle utilisateur proposé est défini comme une ontologie utilisateur générique (figure III.9), comprenant diverses caractéristiques d'un utilisateur, à base de concepts, sous concepts et relations entre les différents concepts.

Afin de réaliser un vocabulaire consensuel, l'ontologie utilisateur a été conceptualisée à partir de la spécification « Information Management SystemLearnerInformation Package » (IMS LIP, 2001). IMS LIP est structuré en onze catégories incluant: Identification, But (Goal), QCL(Qualifications, Certifications et Licences), Compétence, Accessibilité, Activité, Hobby (Interest,Affiliation, Sécurité, Clé et Relation) [12].

Ces catégories sont décrites comme des notions abstraites dans l'ontologie .

En bref, Identification contient des attributs qui aident à identifier une personne (nom, adresse, email, et ..) .

Exemple considérons A1 un apprenant décrit ainsi :

<A1, <Nom, "xxx">, <langage, {"français", "anglais"}>, <media, {"texte", "vidéo"}>>.

L'affiliation inclut information sur la description de l'organisation associée à l'utilisateur/l'apprenant.

Le concept Compétence décrit les compétences associées avec la formation formelle ou informelle de l'utilisateur et son expérience de travail [12].

L'Activité contient les activités liées au travail, formation de l'utilisateur.

L'Accessibilité décrit les préférences des utilisateurs, langues, et handicaps. Les informations sur les passetemps et les activités récréatives sont décrites dans le concept Interest (Hobby).

Le concept Goal (figure III.9) contient l'information sur les buts des utilisateurs/apprenants.

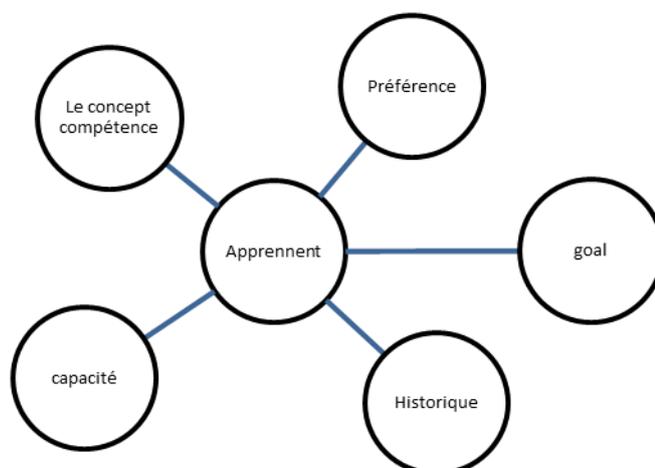


Figure III.9: ontologie apprenant élaboré [12]

L'ontologie utilisateur contient un ensemble de concepts, relations taxonomiques entre les concepts (User, Learner) et des relations non taxonomiques (User works\_onProject).

Ce modèle, représenté graphiquement en Figure.III.10, a été implémenté en utilisant OIModeler et KAON, une boîte à outils pour la gestion d'ontologie et une A.P.I. pour le développement d'applications basées sur les ontologies [13].

La modélisation utilisateur à base d'ontologie nécessite une structure référentielle classique statique (comme les concepts abstraits d'IMSLIP) et une partie adaptative qui doit évoluer à partir de progrès de l'apprenant conformément à son but, domaine d'intérêt, etc. Dans cette perspective on peut dire que l'apprentissage humain rencontre l'apprentissage machine.

Des travaux récents sont menés pour aider à gérer automatiquement l'évolution d'une ontologie.

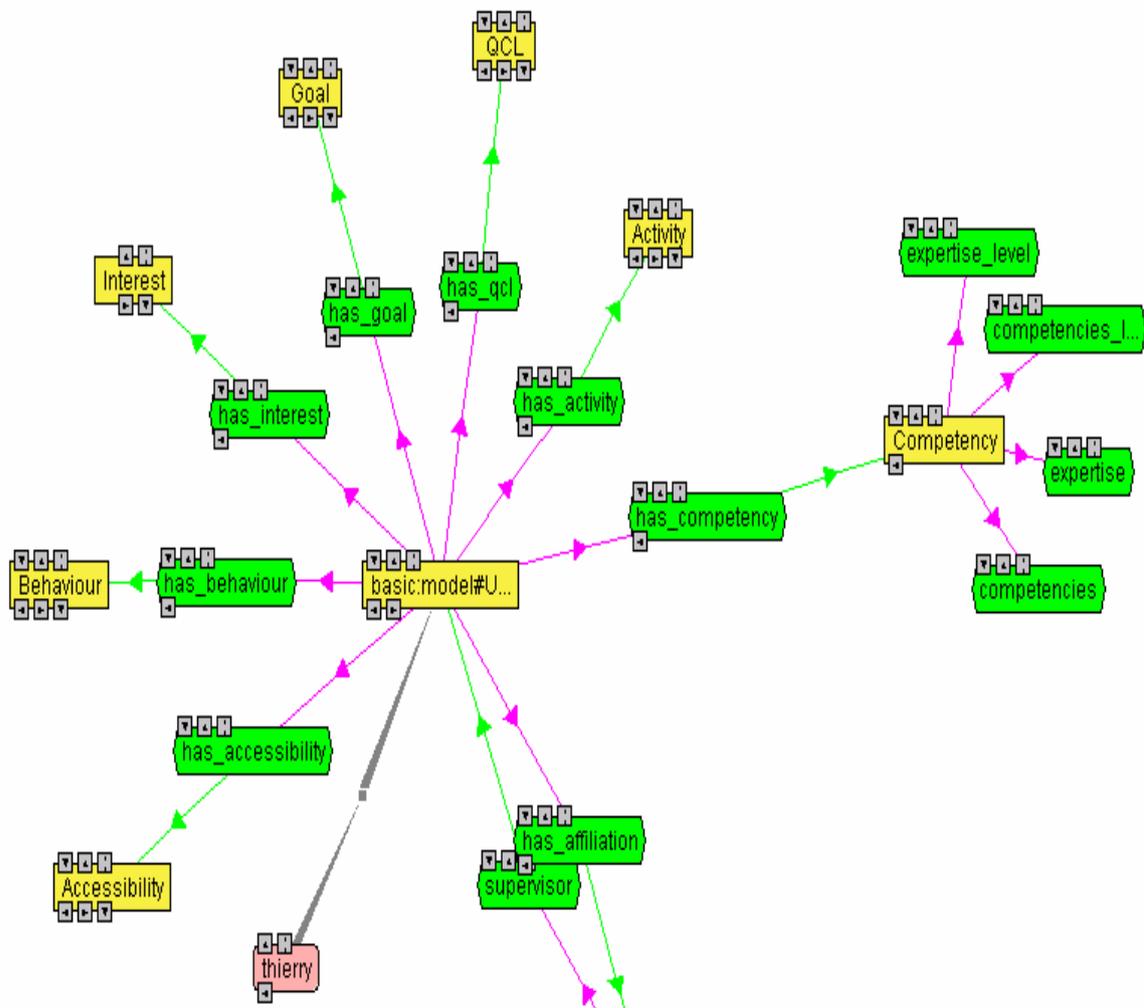


Figure III.10 : L'ontologie utilisateur (concepts abstraits) en OI Modeler [13]

Données de livraison (figure III.11) : Les caractéristiques des utilisateurs peuvent être utilisées pour la personnalisation de services d'enseignement.

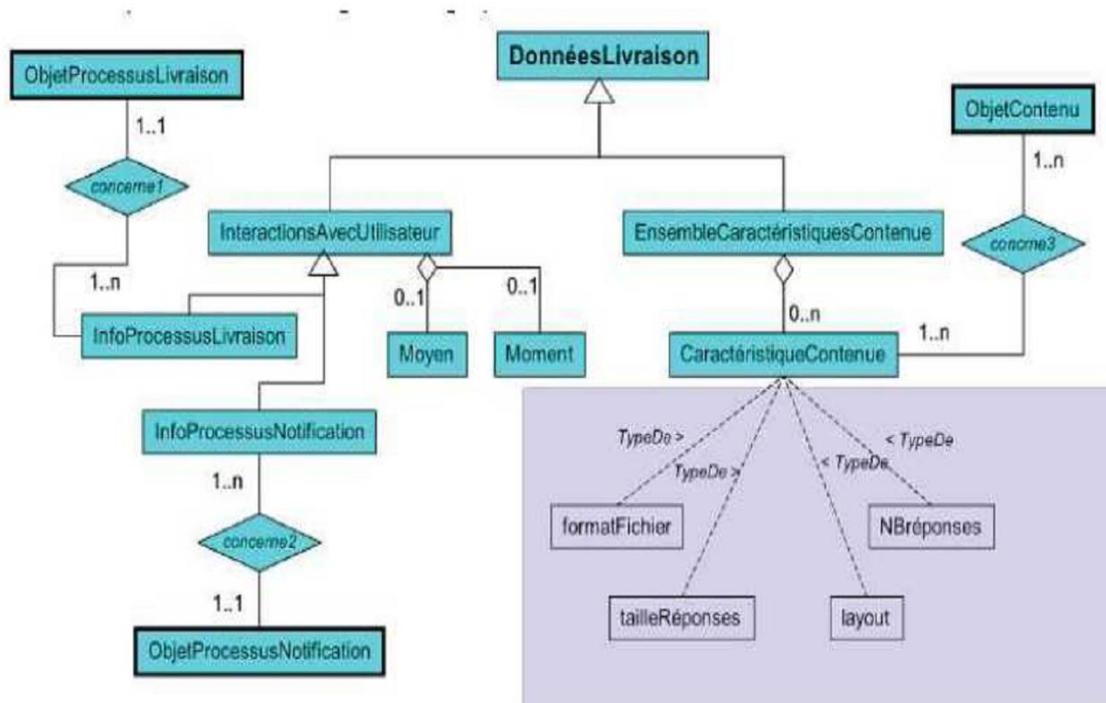


Figure III.11 : la dimension donnée de livraison [14]

Par exemple, en prenant en compte le but de l'apprenant, des agents pédagogiques peuvent proposer différents objets d'apprentissage qui correspondent à l'objectif de l'apprenant et à son niveau d'expertise [14].

### **III.9 Conclusion**

Nous avons abordé dans cette partie le problème de la prise en compte du profil utilisateur dans un système de recherche d'information. La difficulté est due notamment à la modélisation du profil utilisateur et à la conception d'un prototype de système qui met l'utilisateur au centre de la recherche. Cette tâche est encore plus délicate avec la croissance du volume de données.

Nous avons détaillé la notion d'ontologie puisque c'est à base de concept qu'on va modéliser le profil utilisateur.

Le prochain chapitre sera consacré à la conception et réalisation.

## Chapitre IV

### Conception et réalisation

## **Partie 1 : La conception**

### **IV .1 .Introduction**

Après avoir donné une vision générale sur les systèmes de recherches d'information et le profil utilisateur et après avoir donné une vision générale sur les différentes techniques de modélisation du profil utilisateur ,nous allons à présent entamer la partie conception basé sur l'ontologie pour être utiliser dans un chat, ainsi que la conception du système qui va exploiter cette dernière pour permettre aux utilisateurs a interagir entre eux et leur profil.

Nous allons commencer cette partie par une introduction à UML, puis nous présentons ensuite notre architecture du système tout en décrivant le système.

## IV.2.UML

(En l'anglais *Unified Modeling Language*) Langage de modélisation unifié, est un langage de modélisation graphique à base de pictogrammes. Il est utilisé en développement logiciel, et en conception orientée objet. UML est couramment utilisé dans les projets logiciels, UML est un standard En janvier 1997. Parmi les objectifs d'UML : être indépendant des langages de programmation et être adapté à toutes les phases de développement.

### IV.2.1. Les Diagrammes UML

UML propose 14 diagrammes sont dépendants hiérarchiquement et se complètent, de façon à permettre la modélisation d'un projet tout au long de son cycle de vie.

#### IV.2.1.1. Diagrammes structurels ou statiques

Les diagrammes structurels ou statiques (Structure Diagramme) rassemblent :

- Diagramme de classes (Class diagramme)
- Diagramme d'objets (Object diagramme)
- Diagramme de composants (Component diagramme) :
- Diagramme de déploiement (Deployment diagramme) :
- Diagramme des paquetages (Package diagramme)
- Diagramme de structure composite (Composite Structure Diagramme) :
- Diagramme de profils (Profile diagramme (en))

#### IV.2.1.2. Diagrammes comportementaux

Les diagrammes comportementaux (Behavior Diagramme) rassemblent :

- Diagramme des cas d'utilisation (use-cases ou Use Case Diagramme) :
- Diagramme états-transitions (State Machine Diagramme) :

#### IV.2.1.3. Diagrammes d'interaction ou dynamiques

Les diagrammes d'interaction ou dynamiques (Interaction Diagramme) rassemblent :

- Diagramme de séquence (Séquence Diagramme)
- Diagramme de communication (Communication Diagramme)
- Diagramme global d'interaction (Interaction OverviewDiagram)
- Diagramme de temps (Timing Diagramme)

### IV.3. Un réseau social

Désigne un ensemble de personnes réunies par un lien social. À la fin des années 1990, des réseaux sociaux sont apparus sur Internet, réunissant des personnes via des services d'échanges personnalisés, chacun pouvant décider de lire les messages de tel ou tel autre utilisateur.

Ex Facebook, créé en 2004, est le plus connu d'entre eux, et le plus utilisé à ce jour [15].

### IV.4. Conception

#### IV.4.1. Conception l'ontologie

Comme on a dit dans le troisième chapitre Les techniques de modélisation du profil propose dans la technique base sur l'ontologie trois types d'ontologies.

Nous avons choisi la méthode proposée par l'université de Sanford (Les ontologies du domaine) parce qu'elle comporte des étapes claires, simples et faciles à comprendre. Ajouter à cela le fait que l'outil avec lequel nous allons construire l'ontologie en l'occurrence « protégé » est développé par la même université.

**Etape 1** : Déterminer le domaine et la portée de l'ontologie.

- Notre ontologie «chat » va couvrir principalement les domaines d'intérêts avec les profils et les postes des utilisateurs
- Le but de l'utilisation de l'ontologie est de regrouper le différent profil et postes des utilisateurs ainsi que déterminer les différents domaines d'intérêts des utilisateurs avec la liaison entre eux et avec le degré dans laquelle est l'utilisateur intéressé par eux.
- L'ontologie sera utilisée et maintenue par les administrateurs du réseau social

**Etape 2** : Réutiliser des ontologies existantes

Nous avons cherchées pour trouver une ontologie qui regroupe les différents domaines

d'intérêts de l'utilisateur , malgré que on a pas trouver les ontologies requis mais on a trouvé un modèle de domaines d'intérêt proposé par thomas plotkowiak (est un chercheur dans l'institut MCM (Media and Communication Mangement) dans l'université de Saint-Gall a la suisse, sa thèse de doctorat était sur les médias sociaux , il concentrer sur l'influence des réseaux sociaux comme facebook et twitter sur la diffusion de l'information) , Nous allons donc construire une partie de notre ontologie a la base de ce modèle , le reste de l'ontologie est construit à partir de fonctionnalités requise par le système comme les classes profile et postes.

**Etape 3** : Enumérer les termes important de l'ontologie

Les termes importants de notre ontologie est énumérer dans le modèle proposé par thomas plotkowiak , par exemple: education, health, science, economy...etc.

**Etape 4** : définir les classes et la hiérarchie des classes.

Concepts	Description
Class Securityextend Class Computer and internet	Intéressé dans divers domaines de la sécurité (protection, pénétration)
Profile	Regroupe tous les informations
Groupe	Regroupe tous les informations de membre de groupe et lui poste
Message	Contenant le texte du message et l'auteur
..... .....	..... .....

Tableau IV.1 : Les classes et hiérarchie des classes de l'ontologie.

Propriété	Description	concept	type	cardinalité
Nom	Le nom de concept	Touts les concepts	chaine	1,1
Description	La description de concept	Touts les concepts	chaine	1,1
.....	.....	.....	.....	.....

Tableau IV.2 : Les propriétés du modèle de l'ontologie

**Description des relations :**

Dans le langage OWL on n'a pas la possibilité de définir les attribues values, donc on a défini quelque relation comme une classe.

Par exemple on a besoin de définir une relation liaison entre deux domaines avec le poids sur 1 (par exemple 0.3) pour cette relation , donc on a défini une classe liaison qui a comme propriétés « avoir source (le premier domaine de la relation)» et « avoir destination (le deuxième domaine de la relation)» et « poids (le degré de liaison entre le deux domaines)».

Relation	Domaine	Domaine	Description
Degrée de intéresse	Profil	Domaines	De quelle degré L'utilisateur est intéressé par un domine
-Adminegroupe -membergroupe	Profil	Groupe	L'utilisateur peut être membre d'un groupe ou de plusieurs ou un membre de
TableMessage	Profil	Message	L'utilisateur peut venir à lui un message ou plusieurs
.....	.....	.....	.....

Tableau IV.3 : Les Relations entre concepts du modèle de l'ontologie

Etape 7 : créer les instances des classes dans la hiérarchie

Classe	instances
Finance and investement	Accouting Banking .....
Programming and developement	Ada Basic .....
Science	Mathematics .....
.....	.....

Tableau IV.4 : Des exemples d'instances des classes du modèle d'ontologie.

**Le diagramme de classe de l'ontologie :**

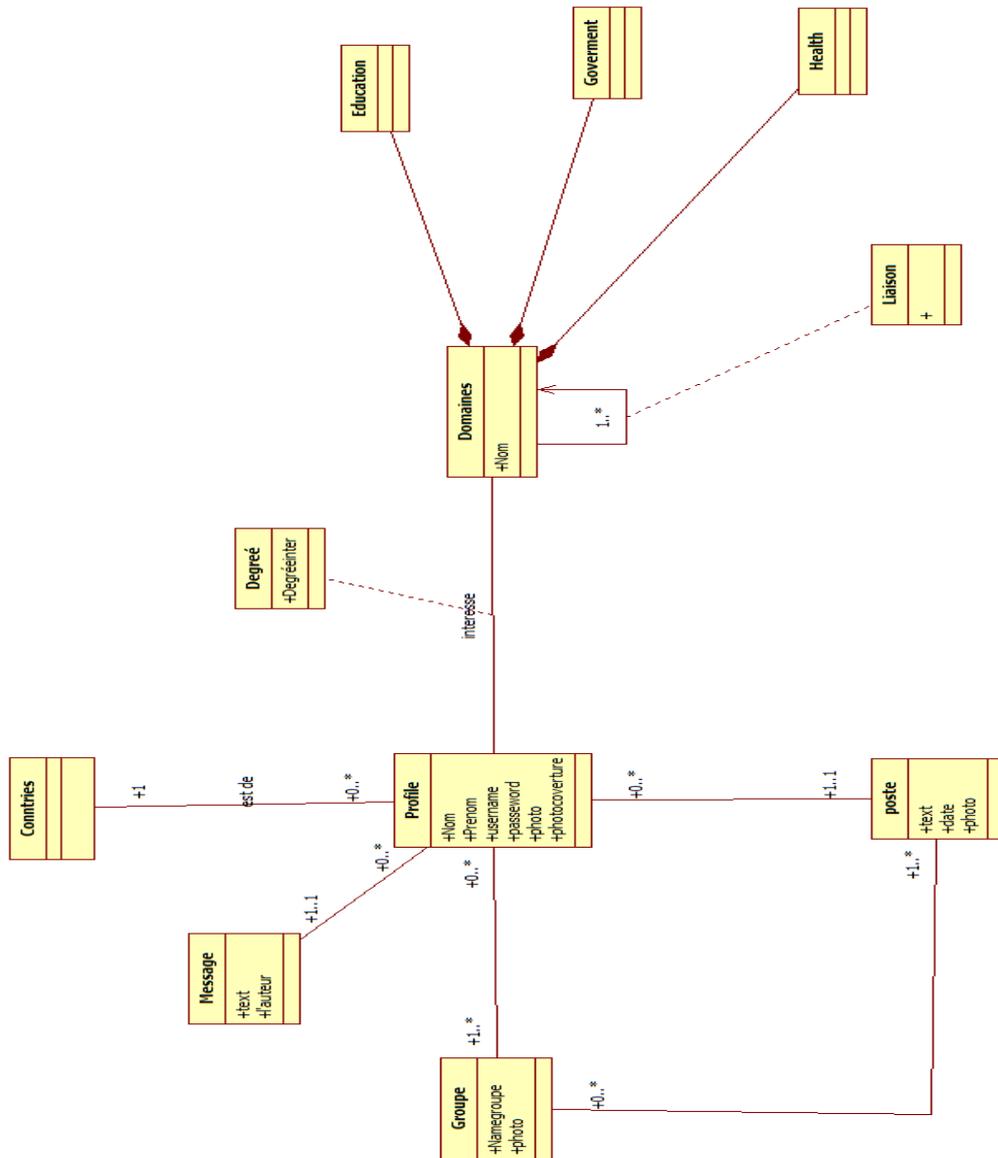


Figure IV.1 : Le diagramme de classe de l'ontologie

### Le Diagramme de classe profil danse application :



Figure IV.2: Le Diagramme de classe profile danse application

Lorsque la connexion est chargée du contenu de classe de l'ontologie à la classe profile danse application

#### IV.4.2. Architecture générale du système

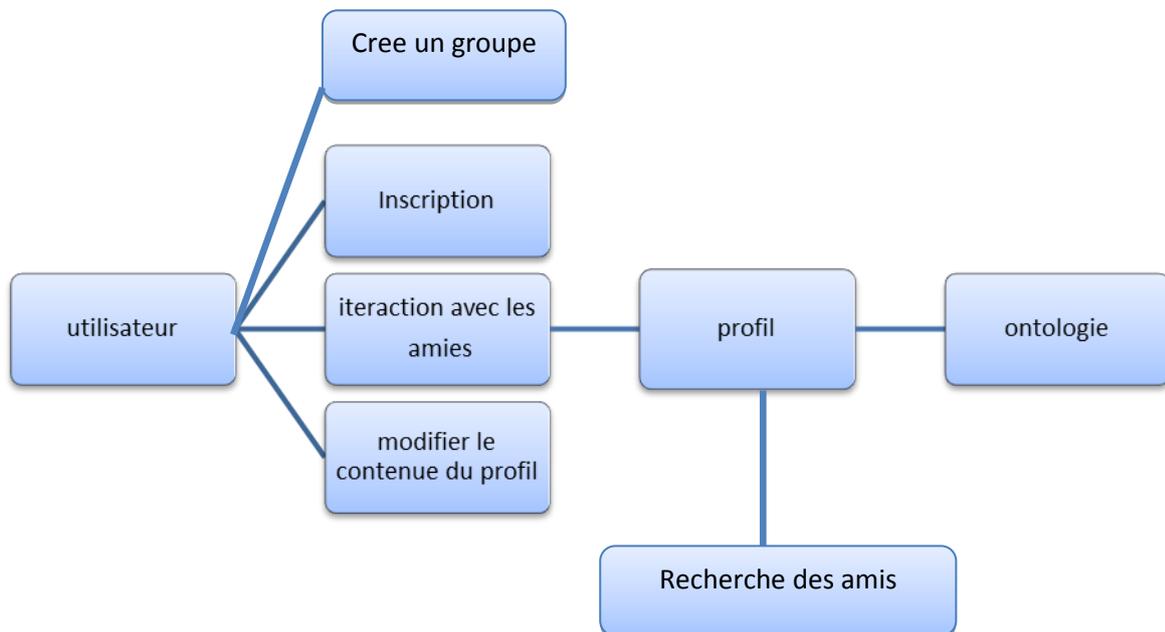


Figure IV. 3: L'architecture générale du système

## IV.5. Description du système

Pour l'exploitation de l'ontologie, notre système se compose de quatre modules principaux :

### IV.5.1. Module d'inscription et créer un groupe

Ce module permet à l'utilisateur qui n'a pas un compte dans System a faire une inscription qui passe par 2 étape.

la première étape : l'utilisateur va passer le test de CAPTCHA

La deuxième étape : l'utilisateur va introduire ses coordonnées (nom, prénom, username , motede passe , country).

Le module d'inscription est un facteur très important dans la performance de système Il contribue à la capacité du système à trouver des personnes que vous connaissez et mettre les informations correctes dans le profil de l'utilisateur.

Ce module permet à l'utilisateur qui crée groupe si a un compte dans System qui passe par 1 étape. : l'utilisateur va introduire ses coordonnées ((nomgroupe , username)

### IV.5.2. Module de recherche des amis

Ce module permet au système de faire une recherche des amis dans un mode automatisé selon : le pays , la liaison familiale ,l'âge , le sexe et le domaine d'intérêt ,ces derniers informations sont collecté lors de la phase d'inscription et peut aussi effectuer la recherche selon les cinq filtres à la fois.

A part de la recherche selon les trois filtres précédents, l'utilisateur peut rechercher dans la liste d' « amis des amis » ainsi qu'une recherche par nom sur le réseau.

### IV.5.3. Module d'interaction avec les amis

Ce module permet à l'utilisateur d'interagir avec ses amis, l'utilisateur peut poster des postes (sans ou avec une image) dans son propre profil, il peut aussi par envoi les message (sans ou avec une image) ,et consulter le profil des amis, envoyer et recevoir les demandes d'amitiés.

### IV.5.4. Module de configuration

Ce module permet à l'utilisateur de modifier ces informations : les domaines d'intérêts, les liaisons familiales, les informations personnelles (nom, prénom, pays,) et la photo de profil et de couverture

### IV.6. Le Diagramme de cas d'utilisation

Les acteurs : il y a deux acteurs

1. L'utilisateur : qui fait l'inscription au début (inclut 5 étapes), qui peut modifier ses informations (les domaines d'intérêts, les liaisons familiales...)
2. Système : qui fait chercher des amis selon les critères proposés, puis interagir avec ses amis

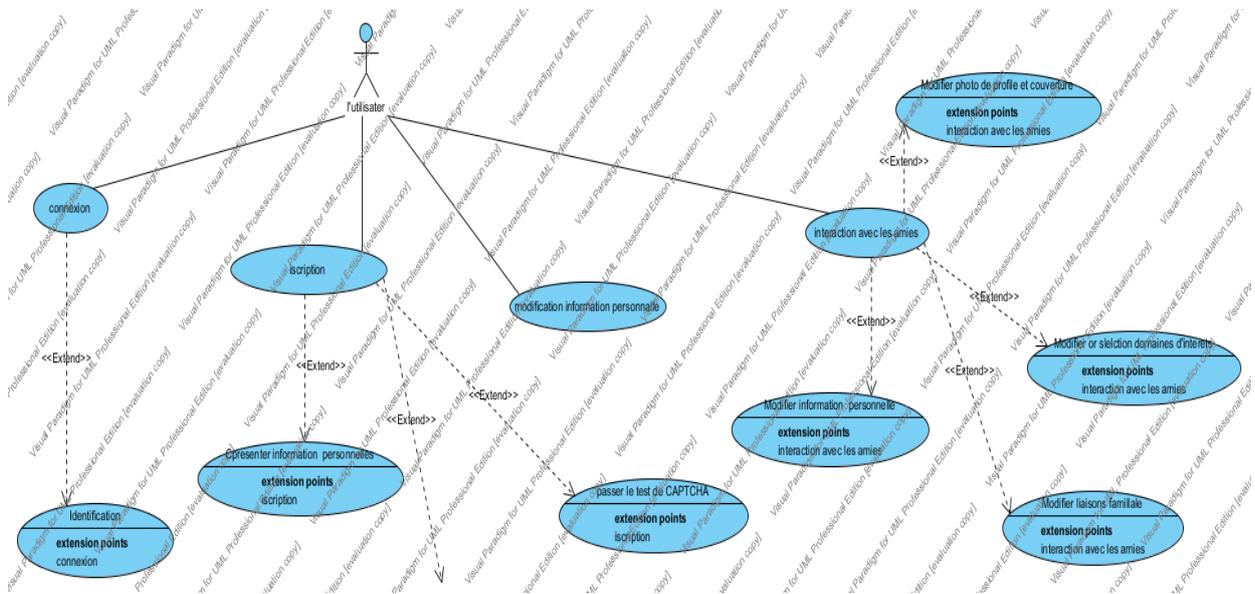


Figure IV.4 : Diagramme de cas d'utilisation de utilisateur [15]

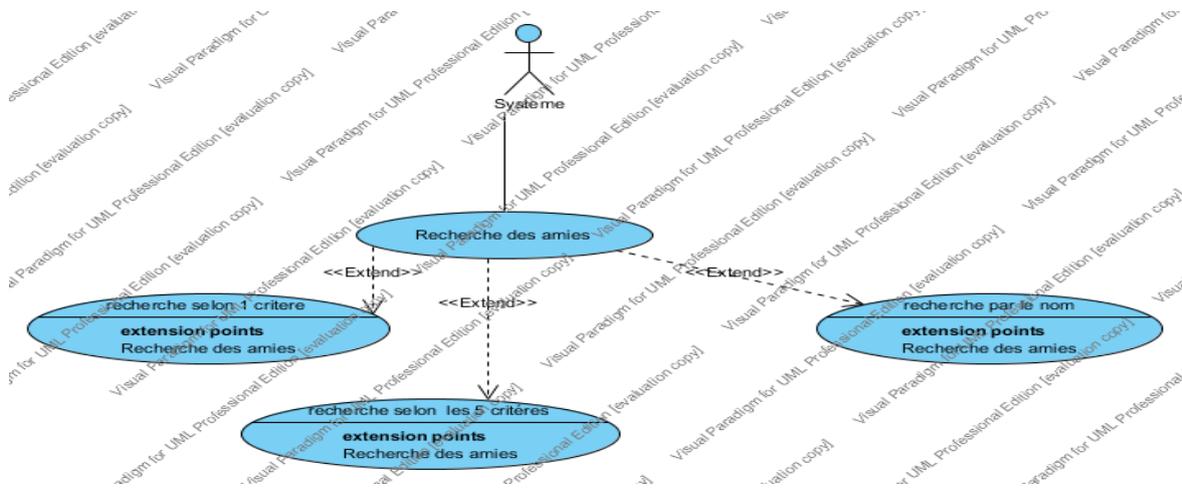


Figure IV.5: Diagramme de cas d'utilisation de Système [15]

**Identification des scénarios :**

L'acteur	Le cas d'utilisation	Le scénario
L'utilisateur	Connexion	S0 : authentification
	Inscription	S0 : Remplissez les cases de renseignements personnels
		S1 : passer le test de CAPTCHA
		S2 : présenter les liaisons Familiales
		S3 : sélectionné une photo de profile et de couverture
	Interaction avec les amis	S0 :consulter le profile pas amies ou amies
		S1 :envoyer les demandes d'amitiés
		S2 : recevoir (accepter ou refuser) les demandes

		d'amitiés
		S3 : envoyer et recevoir les lettre personnels avec ou sans image
		S4 : partager un poste
		S5 : attribuer une mention «j'aime» ou «je n'aime pas» à une poste
	Modification des informations	S0: Modifier informations personnelles
		S1 : Modifier les liaisons Familiales
		S2 : Modifier la photo de profile & couverture
		S3 : Modifier les domines d'intérêts
System	Recherche des amis	S0 : recherche selon le domaine d'intérêt
		S1 : recherche selon la liaison familiale
		S2 : recherche par sexe
		S3 : recherche par nom
		S4 : recherche par nombre d'amis abonnés

Tableau IV.5 : Identification des scénarios

## IV.7. Les diagrammes de séquences

S0 : Lance application de system

S1: Authentification

### Déroulement de l'opération :

Début

L'utilisateur demande l'accès (tapie le nom usage et mot de passe)

Le système vérifie l'existence du l'utilisateur

Si (l'utilisateur existe) le système lui donne l'accès

Sinon il refuse l'accès

Fin si

Fin

### Le diagramme de séquence :

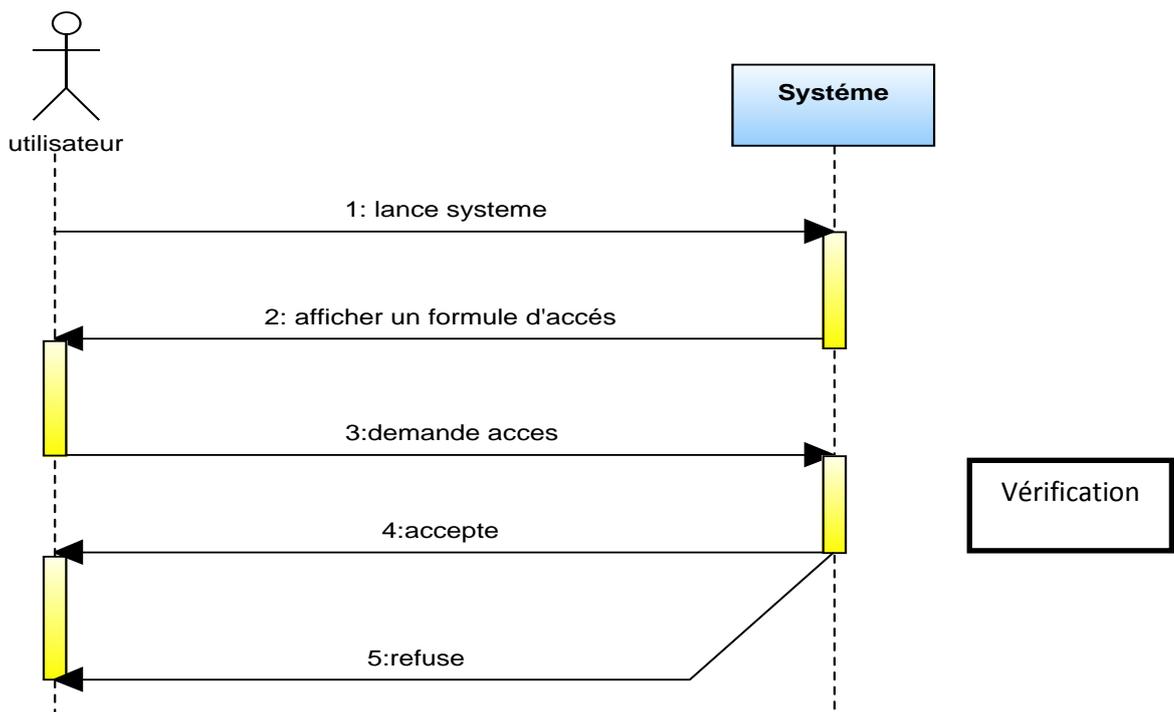


Figure IV.6 : Le digramme de séquence pour l'authentification

S2 : recherche d'amies

**Déroulement de l'opération :**

Début

Le système affiche une liste individuelle pas amie

L'utilisateur sélectionner un amie et demande de envoyer les demandes d'amitiés

Fin

**Le diagramme de séquence :**

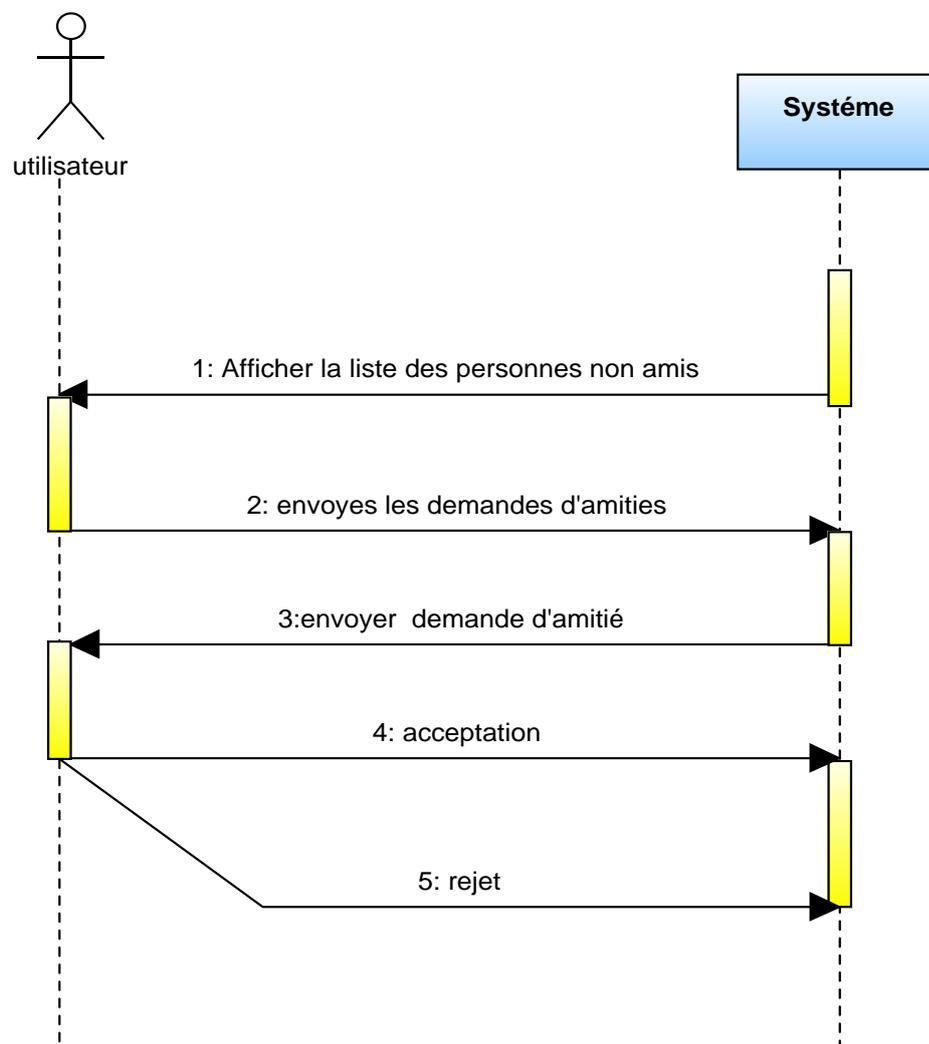


Figure IV.7 : Le diagramme de séquence pour la recherche

## S3 : Modification

**Déroulement de l'opération :**

Début

L'utilisateur demande de modification

Le système affiche un formulaire

L'utilisateur remplir le champ affiché

Systeme est le processus de modification

Fin

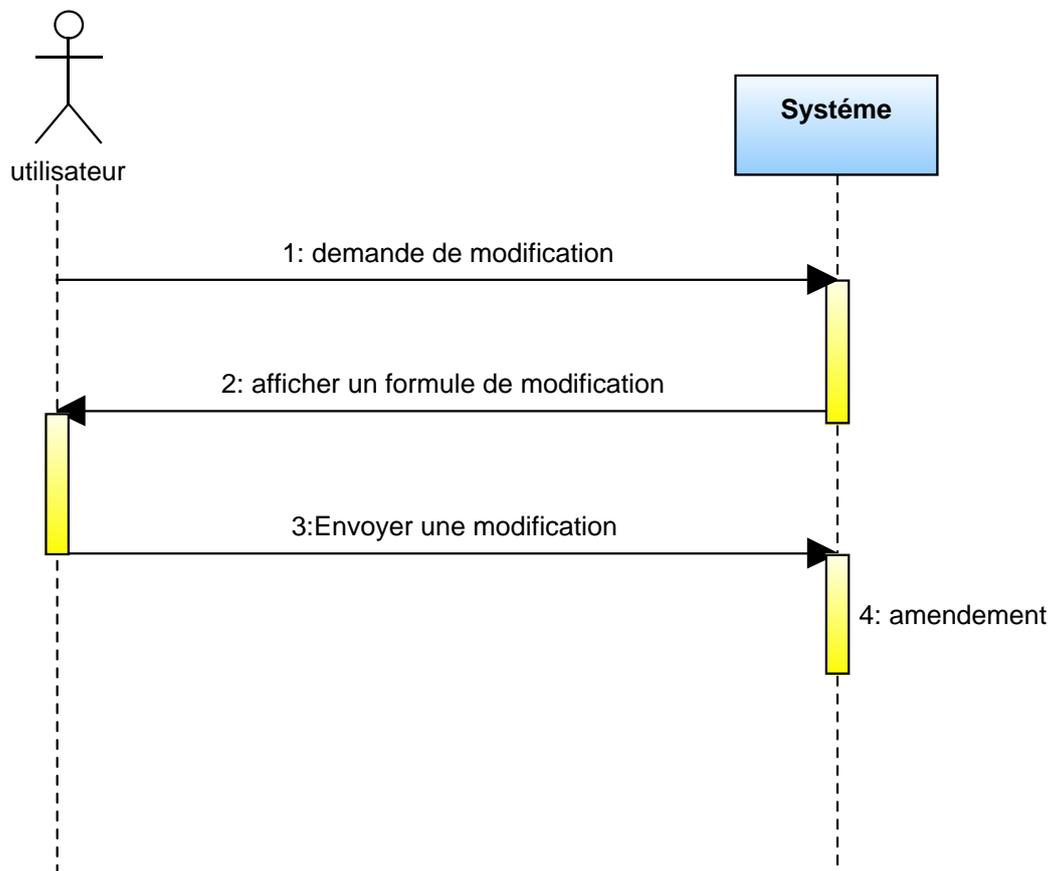
**Le diagramme de séquence**

Figure IV.8 : Le digramme de séquence pour la modification

## **IV.8. Conclusion**

Nous avons présenté dans cette partie le cadre conceptuel de notre système et profils des utilisateurs et son scénario, ainsi construction de ontologie de système qui regroupe les différents domines d'intérêts

## Partie 2 : La réalisation

### IV.9. Introduction

Après avoir dans la première partie de la conception du système et de ses composants et scénarios possible et profil de l'utilisateur en utilisant la conception de l'ontologie.

Nous présenterons l'ensemble des outils nécessaires pour le développement de notre application, et nous détaillerons le processus d'implémentation d'un chat, ainsi que les principales interfaces qui le composent à travers des fenêtres de capture.

### IV.10. Environnement de développement

#### IV.10.1. Protégé

Protégé est un système auteur pour la création d'ontologies. Il a été créé à l'université Stanford et est très populaire dans le domaine du Web sémantique et au niveau de la recherche en informatique.

Protégé est développé en Java. Il est gratuit et son code source est publié sous une Protégé peut lire et sauvegarder des ontologies dans la plupart des formats d'ontologies : RDF, RDFS, OWL, SWRL etc.

Nous utilisons la version 4.3.3.

#### IV.10.2. Le langage SWRL

La sémantique de la règle de Web Langage ( SWRL ) est un langage proposé pour le Web sémantique qui peut être utilisé pour exprimer des règles ainsi que la logique , combinant OWL DL ou OWL Lite avec un sous-ensemble de la Règle Markup Langage (lui-même un sous-ensemble de Datalog ) .

Le cahier des charges a été présenté en mai 2004 au W3C par le Conseil national de recherches du Canada , Network Inférence ( acquis depuis par webMethods ) , et l'Université de Stanford , en association avec l'agent Comité mixte US / UE ad hoc Markup Langage . La spécification a été basé sur une proposition antérieure d'un langage de règles OWL .

Les règles sont de la forme d'une implication entre un antécédent (corps) et par conséquent (la tête ) .Le sens voulu peut être lu comme : chaque fois que les conditions spécifiées dans la soute de l'antécédent ,les conditions spécifiées dans le conséquent doivent également tenir .

### **IV.10.3. Le langage 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 (cofondateur de Sun Microsystems en 1982), présenté officiellement le 23 mai 1995 au SunWorld.

La société Sun a été ensuite rachetée en 2009 par la société Oracle qui détient et maintient désormais Java.

La particularité et l'objectif central de Java est que les logiciels écrits dans ce langage doivent être très facilement portables sur plusieurs systèmes d'exploitation tels que UNIX, Windows .

### **IV.10.4. NetBeans**

NetBeans est un environnement de développement intégré (EDI), placé en open source par Sun en juin 2000 sous licence CDDL et GPLv2 (Common Development and Distribution License). En plus de Java, NetBeans permet également de supporter différents autres langages, comme Python, C, C++, JavaScript, XML, Ruby, PHP 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).

Conçu en Java, NetBeans est disponible sous Windows, Linux, ou sous une version indépendante des systèmes d'exploitation (requérant une machine virtuelle Java). Un environnement Java Development Kit JDK est requis pour les développements en 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.

## IV.11. Scénarios d'exécution avec captures d'écrans

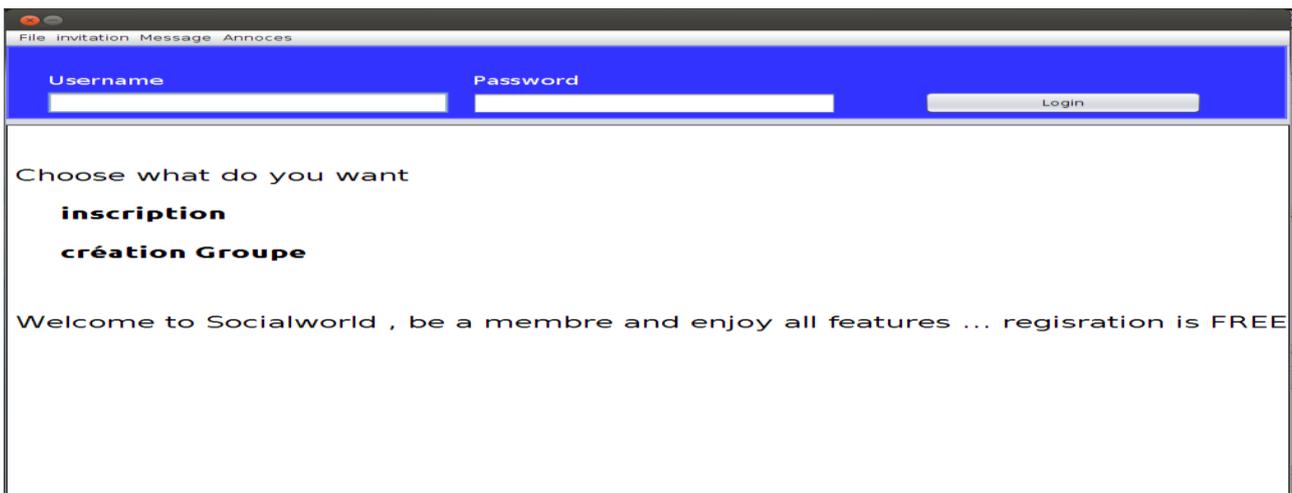


Figure IV.9 : l'interface d'application

### ❖ Inscription

a) le test de CAPTCHA :

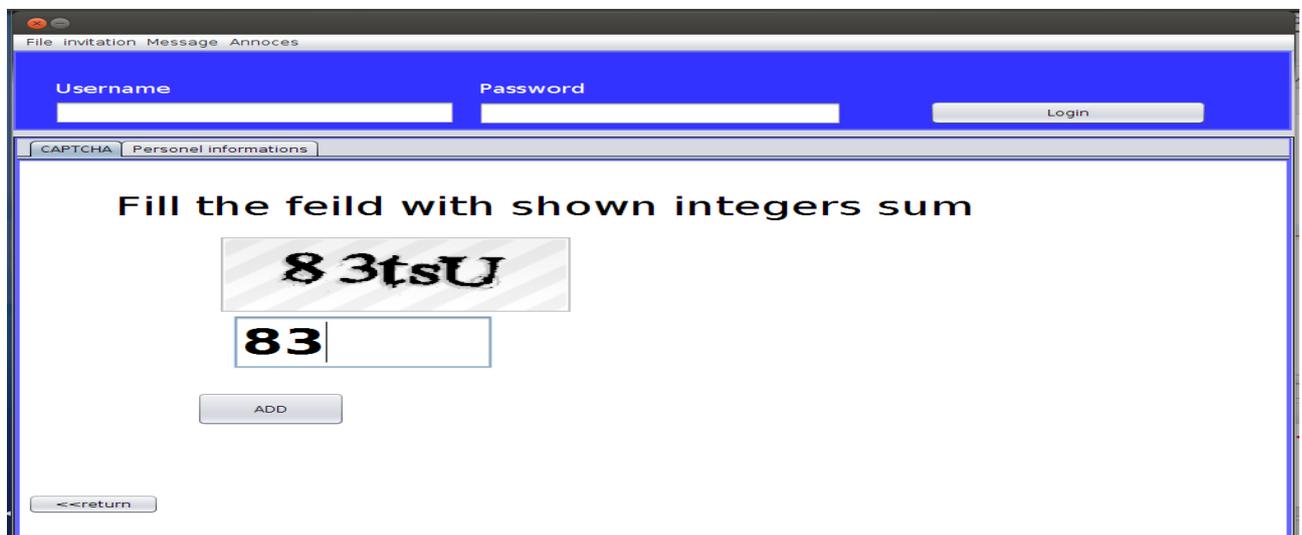


Figure IV.10 : le test de CAPTCHA

**b) Présentation des informations personnelles :**

File invitation Message Annoces

Username Password Login

CAPTCHA Personel informations

Private information private

First name fares

Family name fares

Username fares

Password \*\*\*\*\*

Sex  Female  Male

date of birth 1 / 23 / 1990

country algeria

<<return Welcome to your world

Figure IV.11 : Présentation des informations personnelles

**❖ Créer un groupe**

File invitation Message Annoces

Username Password Login

groupe informations

Welcome to Socialworld , if want to create a private groupe

name of Groupe JaVA

adminname mourad

ok

<< return

Figure IV.12:créer un groupe

❖ Le compte (Home) obtenue :

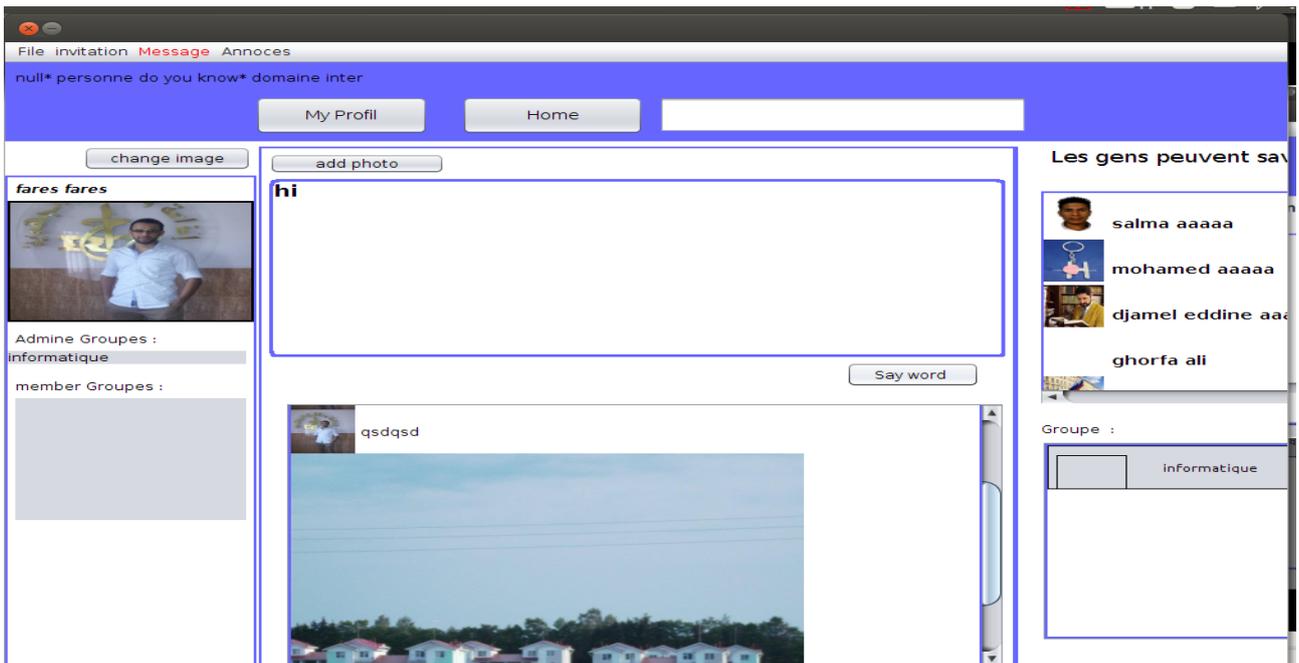


Figure IV.13: la page Home

❖ Le Profil obtenue



Figure IV.14: le profil

Click Personale informations pour obtenue

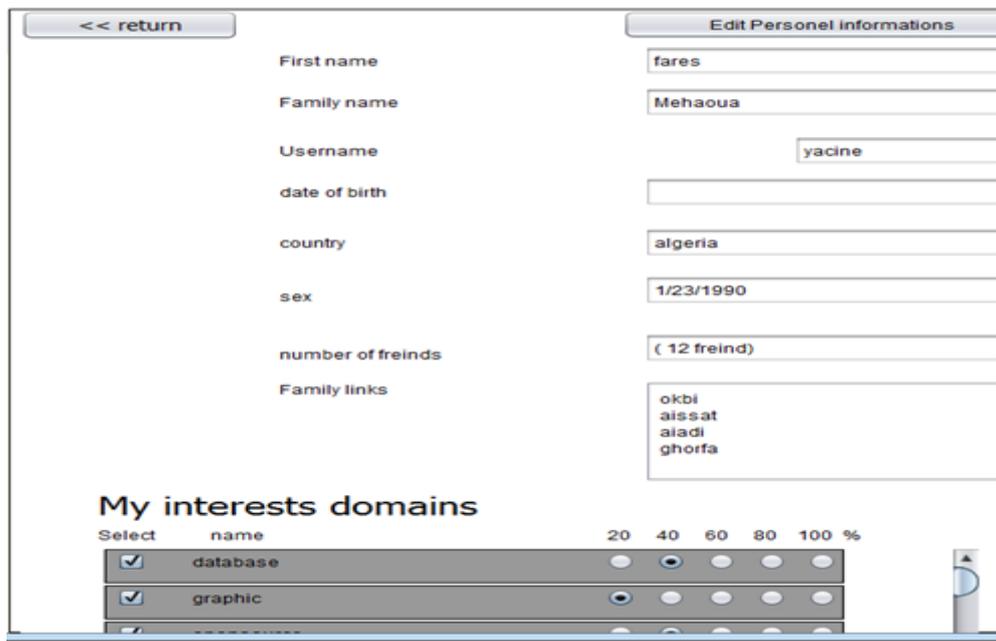


Figure IV.15: information personnelle

❖ Recherche des amis :

Le système recherche automatiquement en fonction de l'information fournie par l'articulation

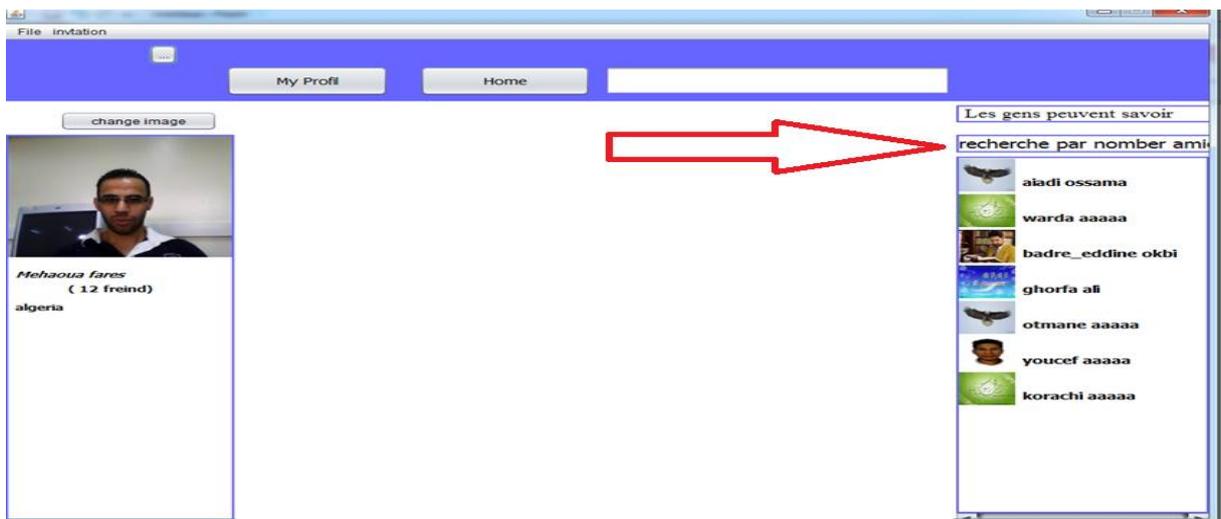


Figure IV.16: Les personnes trouvées

## ❖ Changer la photo de couverture et de profil

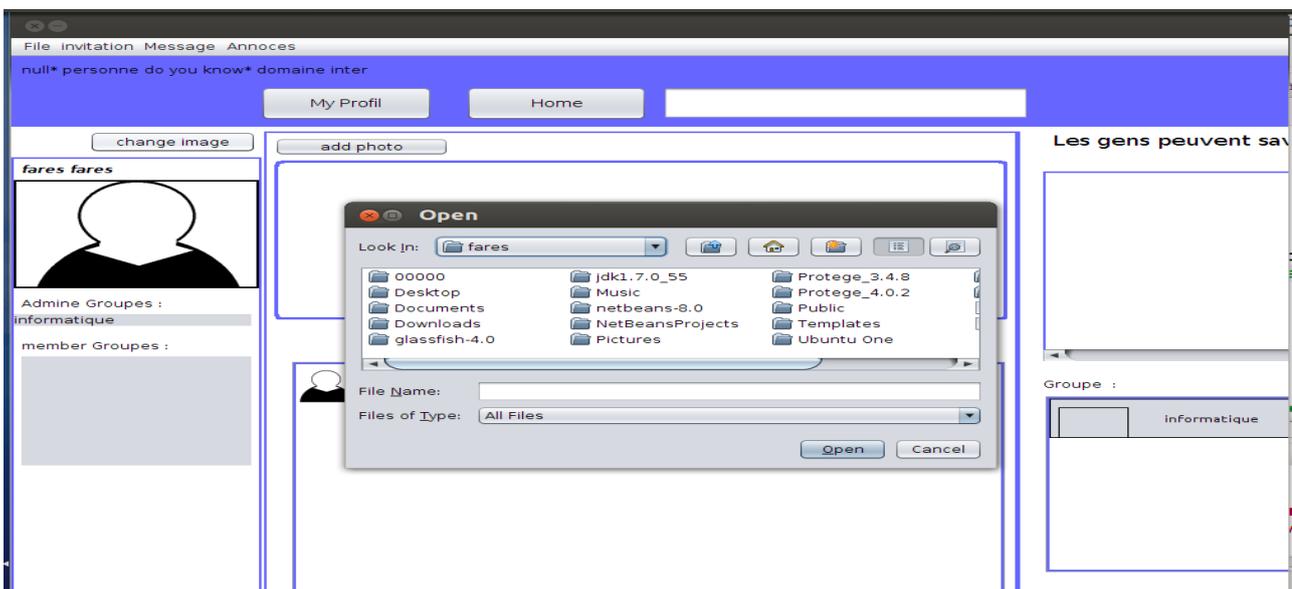
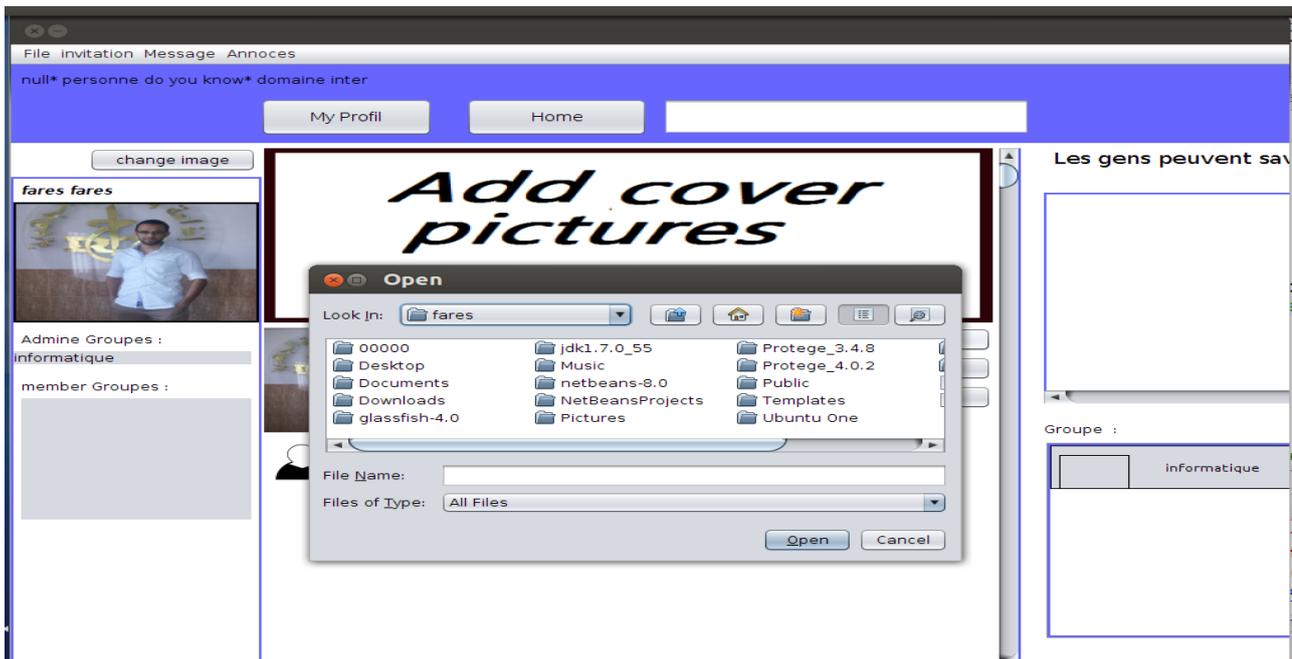


Figure IV.17: Ajouter la photo de couverture et de profil

## ❖ Interaction entre utilisateur

### Dans un Groupe

#### 1 - Mettre le texte ou l'image

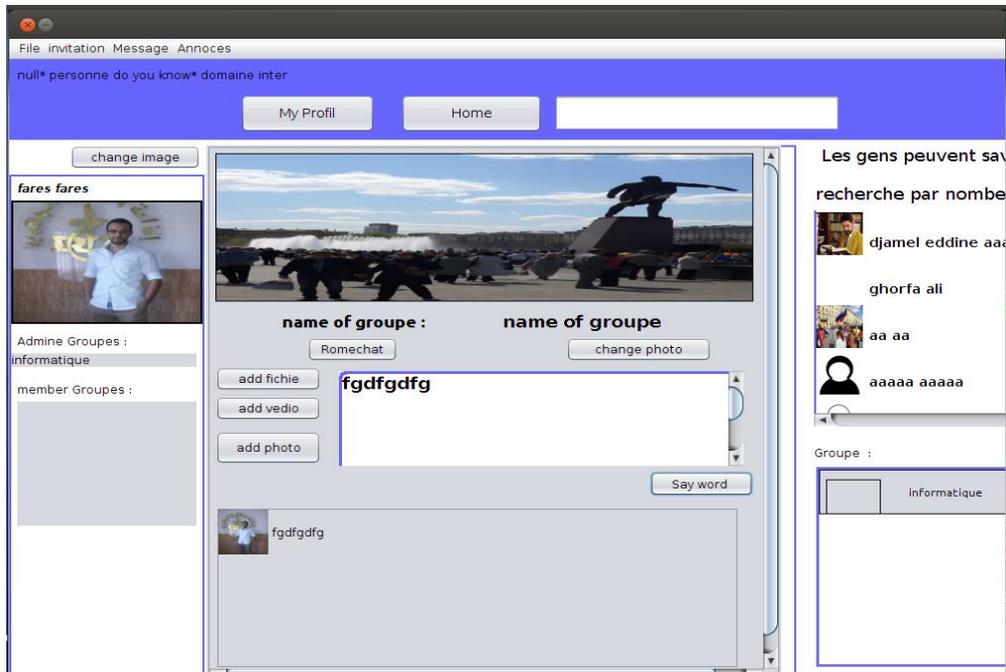


Figure IV.18: la page groupe

#### 2- « Room chat »

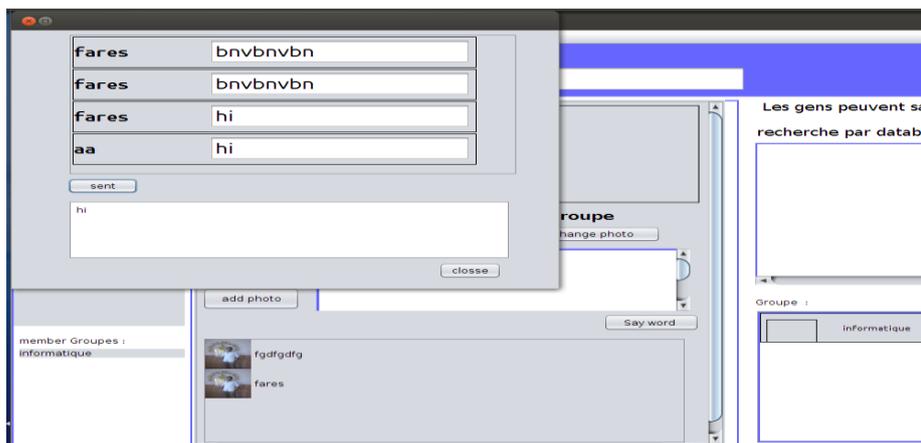


Figure IV.19: « Room chat » dans un groupe

### ❖ Interaction par Message

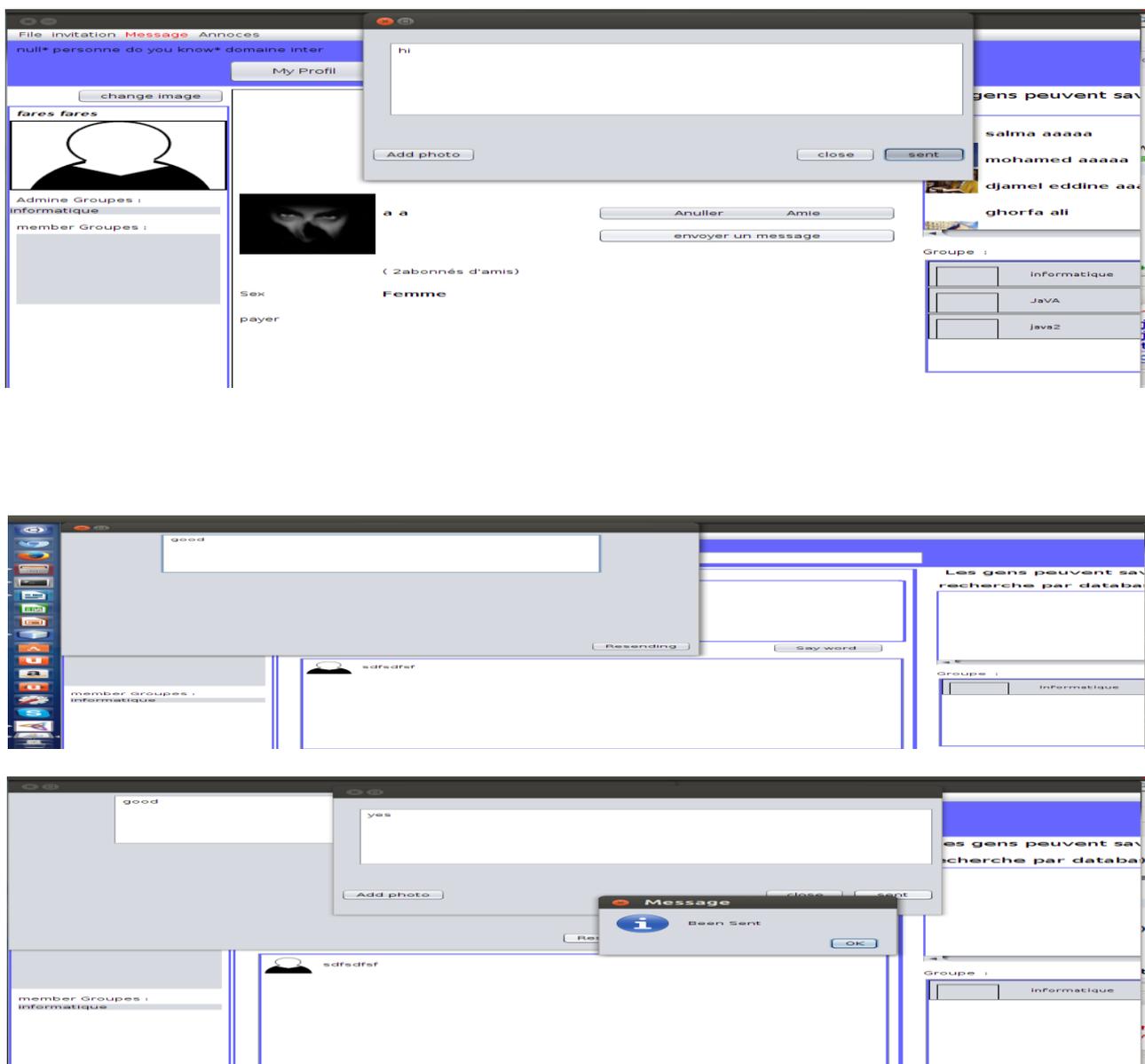


Figure IV.20: interaction par Message

## **IV.12. Conclusion**

Dans cette dernière partie, nous avons présenté la réalisation de notre système ce dernier a été commencé par l'exploration des outils utilisés, ensuite la description du système à travers les captures d'écran.

## **Conclusion générale et perspective**

Dans ce projet, nous avons utilisé le profil utilisateur et principalement son domaine d'intérêt pour filtrer les données

La modélisation du profil a été basée sur l'ontologie qui a donné beaucoup de richesse à la recherche d'information

Notre étude de cas a été réalisée au profit d'un chat avec lequel l'utilisateur accède au réseau et entamer la conversation (communication) avec les utilisateurs qui ont le même profil

Nous pouvons envisager comme perspectives :

- Enficher le profil utilisateur avec d'autre information telles que : qualité, données de livraison, données de sécurité ...etc.
- Utiliser autre technique autre que l'ontologie pour modéliser le profil exemple : les techniques statiques.
- Généraliser la modélisation sur plusieurs cas.
- De mettre l'application sous web.

## Bibliographie

- [1] Jean-Pierre Chevallet, Philippe Mulhem, Laurence Nigay «Conception des Systèmes de Recherche d'Information Multimédia Gilles Gauthier»S
- [2] Hamdi Chaker, Max Chevalier, Chantal Soulé-Dupuy, André Tricot «Système de recherche d'information pour les tâches métier»
- [3] Asma Hedia Brini «Un Modelé de Recherche d'Information basé sur les Réseaux Possibilistes»
- [4] Rezeg Khaled «Découverte des services dans les systèmes d'information géographiques réparti sous réseau ad-hoc» thèse DOCTORAT EN SCIENCES EN INFORMATIQUE, 2010/2011
- [5] Romaric Besançon «Cours Master Recherche Paris 13 Recherche et extraction d'information»
- [6] Laurianne Sitbon, Patrice Bellot, Philippe Blache «Vers une recherche d'information adaptée aux utilisateurs dyslexiques»
- [7] Hatem Hadad «Extraction et Impact des connaissances sur les performances des Systèmes de Recherche d'information» thèse pour obtenir le grade de DOCTEUR DE L' UNIVERSIT OSEPH FOURIER Mohamed, 2007
- [8] Dimitre Kostadinov «Personnalisation de l'information : une approche de gestion de profils et de reformulation de requêtes» UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN-YVELINES
- [9] «profilutilisateur»[http://fr.wikipedia.org/wiki/Profil\\_utilisateur#Contenu\\_d.27un\\_profil](http://fr.wikipedia.org/wiki/Profil_utilisateur#Contenu_d.27un_profil),  
visité en mars 2014
- [10] Zemirli W.Nesrine «Vers le développement d'un système de recherche d'information personnalisé intégrant le profil utilisateur» 2003/2004
- [11] Lynda Tamine-Lechani «Approche statistique pour la définition du profil d'un utilisateur de système de recherche d'information.» Institut de Recherche en Informatique de Toulouse Equipe Systèmes d'Information Généralisés
- [12] Wahiba Nesrine Zemirli « Modèle d'accès personnalisé à l'information basé sur les Diagrammes d'Influence intégrant un profil utilisateur évolutif » l'Université Paul Sabatier de Toulouse III, juin 2008
- [13] Amokrane Belloui « L'usage des concepts du web sémantique dans le filtrage d'information collaboratif » Mémoire présenté pour l'obtention du diplôme de Magistère, Institut National d'Informatique Alger, 2008
- [14] Khalida Farida Khelili « Modélisation à base d'ontologie d'une plate-forme de E-Learning » Mémoire En vue de l'obtention du diplôme de Magistère, Université Kasdi Merbah de Ouargla, 28 mai 2009

[15] Ghorfa Yacine, Aiadi Oussama « Utilisation des technologies du web sémantique dans les réseaux sociaux » Mémoire de fin d'études Pour l'obtention du diplôme de Master en Informatique Université Kasdi Merbah de Ouargla, 2013