

M/004.474

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la Recherche scientifique

Université de Guelma

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Mémoire de fin d'étude Master



Département d'Informatique

Spécialité : Ingénierie du médias

18/862

---

**Thème : Les systèmes de recommandation  
personnalisés en e-learning**

---

Présenté par :

-Maalala Khalid.

-Dridi Salah Eddine.

Sous la direction de :

Madame Benhamdi Soulef.

Juin 2014



# Remerciement

*C'est avec un grand plaisir que nous apportons ce modeste Travail à tous ceux qui nous ont gratifiés de leur soutien et de leur Confiance.*

*Louanges à Dieu, qui nous a donné vie et santé pour le parachèvement de ce travail. Et nous voudrions exprimer toute notre gratitude à :*

*Nos familles, qui durant nos études, nous ont toujours donné la possibilité de faire ce que nous voulions et qui ont toujours cru en nous.*

*Madame Benhamdi Soulef pour la confiance qu'elle nous a témoignée en acceptant de diriger ce travail et pour nous avoir accordé de son temps et avoir mis à notre disposition ses compétences et ses conseils pour une meilleure maîtrise du sujet.*

*Membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail Et de l'enrichir par leurs propositions.*

*Tous les enseignants du département d'Informatique, qui ont assisté à nos débuts, pour leurs précieux conseils*

*Nos collègues de la fin de cycle, qui nous ont donné leurs encouragements toute la durée de réalisation de ce travail.*

*Toutes les personnes qui nous ayant aidé et soutenu de près ou de loin tout le long de ce travail.*

***Khalid & Salah***





# *Dédicaces*

*Louage à « ALLAH » maître de l'univers et paix et salut sur  
notre prophète « Mohamed ».*

*Je dédie ce travail à ma chère mère qui a sacrifiée sa vie pour que  
je puisse réussir et qui par son attention son aide son amour son  
compréhension et son soutien m'a permis d'atteindre mes  
objectifs.*

*Merci et que dieu t'accorde une longue vie.*

*A mon père pour son dévouement qui mon permis d'accéder aux  
études supérieure et faire de moi ce que je suis aujourd'hui.*

*Merci et qui dieu t'accorde une longue vie.*

*A mes frères que dieu les garde en son vaste paradis.*

*A ma sœur .*

*Merci et que dieu t'accorde une longue vie.*

*A toute ma famille.*

*A mes amis pour leur sincérité et leur soutien, leur grande qualité  
humaine, et leurs moments inoubliables qu'on a pu passer  
ensemble.*



# *Dédicaces*

*Louage à « ALLAH » maître de l'univers et paix et salut sur notre prophète « Mohamed ».*

*Je dédie ce mémoire de fin d'études à :*

*Mon très cher père et cher mère, Mes chers frères, et mes chères sœurs pour leur affection, compréhension et patience.*

*A tous mes amis et mes collègues.*

*A tous ceux qui ont une relation de proche ou de loin avec la réalisation du présent mémoire.*

*Khalid*





## **Résumé**

Dans ce mémoire nous nous intéressons à l'utilisation des systèmes de recommandation dans le domaine du e-learning. En effet, nous utilisons une démarche de recommandation qui permet de sélectionner automatiquement les objets (cours) d'apprentissage adaptés aux apprenants. Cela est possible grâce à un classificateur automatique qui permet de décider si un objet d'apprentissage est pertinent ou non pour un apprenant particulier. Ce classificateur est construit par la méthode CSHTRL (Cold Start Hybrid Taxonomique Recommender for e-Learning).

**Mot clés :** e-learning, systèmes de recommandation, CSHTRL.

# Sommaire

<b>Sommaire.....</b>	<b>1</b>
<b>Liste des figures.....</b>	<b>5</b>
<b>Liste des tableaux.....</b>	<b>6</b>
<b>Introduction générale .....</b>	<b>7</b>
<b>Chapitre 1: Le e-learning.....</b>	<b>10</b>
I.1 Introduction.....	10
I.2. Mieux comprendre le e-learning.....	10
I.2.1. Concepts de base du e-learning.....	10
I.2.1.1. E-learning et e-formation.....	11
I.2.1.2. Autres modes de formations.....	11
I.2.2. Les enjeux du e-learning.....	12
I.3. Modélisation des apprenants.....	12
I.3.1. Acquisition des données du modèle d'apprenant.....	13
I.3.1.1. Modèle explicite d'apprenant.....	13
I.3.1.2. Modèle implicite d'apprenant.....	14
I.3.2. Dimensions des informations des apprenants.....	14
I.3.3. Présentation des principaux modèles existants.....	16
I.3.3.1. ANSI TS 130 Student Educational Record.....	16
I.3.3.2. vCard.....	17
I.4.3.3. eduPerson.....	17
I.3.3.4. Educational Modelling Languages (EMLs).....	17
I.3.3.5. Le modèle d'apprenant de Chen et Mizoguchi.....	18
I.3.3.6. IEEE Public and Private Information – Learner specification (PAPI).....	19
I.3.3.7. IMS Learner Information Package (IMS LIP).....	20
I.3.3.8. European Diploma Supplement (EDS).....	23

I.3.3.9. Dolog Learner Profile (Dolog LP).....	23
I.3.3.10. Friend of a Friend (FOAF).....	24
I.4. Modélisation des ressources pédagogiques.....	25
I.4.1. Présentation des principaux modèles existants.....	26
I.4.1.1. Le modèle DublinCore.....	26
I.4.1.2. Le modèle ARIADNE.....	26
I.4.1.3. Le modèle LOM.....	28
I.5. Conclusion.....	30
<b>Chapitre 2 : Les systèmes de recommandation.....</b>	<b>32</b>
II.1. Introduction.....	32
II.2. Définition du filtrage d'information.....	32
II.3. Brève Historique des systèmes du filtrage d'information.....	32
II.4. Domaines d'application du filtrage d'information.....	33
II.5. Caractéristiques d'un système du filtrage d'information.....	33
II.6. Mécanisme du fonctionnement d'un système du filtrage d'information.....	34
II.7. Types des systèmes du filtrage d'information.....	35
II.7.1. Systèmes du filtrage à base de règles.....	36
II.7.2. Limites du filtrage à base de règles.....	36
II.7.3. Systèmes du filtrage par contenu.....	36
II.7.4. Limites du filtrage par contenu.....	37
II.7.5. Systèmes du filtrage collaboratif.....	38
II.7.6. Limites du filtrage collaboratif.....	39
II.7.7. Filtrage hybride.....	40
II.8. Calcul de prédiction.....	40
II.8.1. Prédiction dans le filtrage cognitif.....	40
II.8.2. Prédiction dans le filtrage collaboratif.....	40



II.8.2.1. Algorithmes à base de « mémoire ».....	41
II.8.2.2. Algorithmes d'apprentissage en ligne.....	41
II.8.2.3. Algorithmes à base de « modèle ».....	41
II.9. Une note sur l'évaluation des systèmes de recommandation.....	42
II.10. Conclusion.....	42
<b>Chapitre 3 : Etude conceptuelle.....</b>	<b>45</b>
III.1. Introduction.....	45
III.2. Motivations et Objectif.....	45
III.3. CSHTR (l'origine de CSHTRL).....	46
III.3.1. Objectifs de CSHTR.....	46
III.3.2. Description de la technique.....	46
III.3.2.1. La représentation du monde visé.....	46
III.3.2.2 Les regroupements.....	47
III.3.2.3 Capturer les préférences taxonomiques.....	47
III.3.2.4 Comportement de la technique dans une situation de démarrage à froid.....	47
III.3.2.5 La génération de la recommandation.....	48
III.3.3 Critiques.....	49
III.3.3.1. Avantages.....	49
III.3.3.2 Inconvénients.....	49
III.4. Présentation de CSHTRL.....	49
III.4.1. Description formelle.....	50
III.4.1.1 Profile apprenant.....	50
III.4.1.2 Description de l'environnement.....	51
III.3 Conclusion.....	57
<b>Chapitre 4 : Implémentation.....</b>	<b>59</b>
IV.1. Introduction.....	59

IV.2. Outils et Implémentation.....	59
I.V.2.1. Page apprenant.....	63
I.V.2.2. Page administration.....	63
I.V.2.3. Page enseignant.....	64
IV.3. Expérimentation.....	65
4.3.1. Les résultats de l'expérimentation.....	66
IV.3.2. Posttest.....	67
IV.4. Conclusion.....	67
<b>Conclusion générale.....</b>	<b>68</b>
<b>Références et bibliographie.....</b>	<b>69</b>

## Liste des tableaux

<b>Tableau 4.1</b>	<b>informations des apprenants</b>	<b>Page 65</b>
<b>Tableau 4.2</b>	<b>les évaluations</b>	<b>Page 65</b>
<b>Tableau 4.3</b>	<b>les évaluations de l'apprenant a1</b>	<b>Page 66</b>
<b>Tableau 4.4</b>	<b>les items recommandés à l'apprenant a1</b>	<b>Page 66</b>
<b>Tableau 4.5</b>	<b>résultat post-test</b>	<b>Page 67</b>



## Introduction générale

Chaque jour, les technologies progressent, les métiers évoluent, l'organisation change, les méthodes de management se transforment. Les besoins augmentent tant pour la formation initiale que continue. Mais les budgets disponibles et surtout le temps qu'il est possible de dégager ne sont pas extensibles à l'infini. C'est la raison pour laquelle les outils construits sur l'Internet, qui offrent d'immenses atouts (outre l'économie considérable en temps et en déplacement) émergent à très grande vitesse dans les pays qui ont pris de l'avance (notamment en Amérique du Nord). *[CREF]*

Avec l'avènement des technologies de l'information et de la communication (TIC), nous devons, dès à présent, « penser apprentissage rapide et efficace ». Ces dernières années ont connu un fort essor de la FOAD (Formation Ouverte et À Distance). Cet essor est dû en grande partie au développement des TICE (Technologies de l'Information et de la Communication dans l'Enseignement), fait apparaître un nouveau mode d'apprentissage connu sous le nom de « e-learning ». Celui-ci est basé sur l'accès à des formations en ligne, interactives et parfois personnalisées, diffusées par l'intermédiaire d'un réseau (Internet ou Intranet) ou d'autres média électroniques. Cet accès permet de développer les compétences des apprenants, tout en rendant le processus d'apprentissage indépendant du temps et du lieu. Ainsi, le e-learning est le nom donné actuellement à une phase importante de l'introduction des TIC dans la formation. Il s'agit d'une évolution rapide des technologies pour l'apprentissage, rendue possible par le développement planétaire de l'Internet. Mais, la quantité d'information sur ce dernier ne cesse de croître à chaque moment, alors ceci rend la tâche plus difficile aux utilisateurs de trouver ce qu'ils veulent effectivement. De plus, l'intégration de la technologie dans le processus d'apprentissage nécessite d'aborder la problématique de l'amélioration de qualité d'enseignement et d'apprentissage, plutôt d'être considérée juste comme un nouveau moyen flexible de livraison des ressources pédagogiques [Nichols, 2003].

Notons que le but de la personnalisation est de fournir aux utilisateurs les recommandations qu'ils désirent selon leurs caractéristiques individuelles.

Les systèmes de recommandation sont apparus récemment pour désigner la technologie de recommandation. Ces derniers sont traditionnellement concentrés sur des activités de commerce électronique, pour choisir et suggérer l'achat de différents produits (aux utilisateurs/consommateurs), et maintenant se sont intégrées dans le secteur éducatif (e-Learning) pour ajouter plus d'amélioration.

Notre projet consiste à réaliser une technique qui intègre les systèmes de recommandation en e-learning, cette technique est CSHTML dans but d'apporter une meilleure solution au problème de démarrage à froid.

# Chapitre 1 : Le e-learning

---



## Chapitre 1 : Le e-learning

### I.1 Introduction

Avec l'avènement des technologies de l'information et de la communication (TIC), nous devons, dès à présent, « penser apprentissage rapide et efficace », avec un minimum de problèmes d'organisation, de logistique et surtout de perte de temps. L'e-learning est la solution. C'est le nom donné actuellement à une phase importante de l'introduction des TIC dans la formation. Il s'agit d'une évolution rapide des technologies pour l'apprentissage, rendue possible par le développement planétaire de l'Internet.

### I.2. Mieux comprendre le e-learning

« Comprendre est le commencement d'approuver ». Spinoza

#### I.2.1. Concepts de base du e-learning

##### I.2.1.1. E-learning et e-formation

Le **e-learning** consiste à utiliser les ressources de l'informatique et de l'Internet pour acquérir, à distance, des connaissances. Les services et les contenus de ce type d'apprentissage sont offerts via un ordinateur relié à un intranet, à un extranet ou encore à internet.

En outre, le e-learning est un domaine révolutionnaire, un changement dans le champ de l'apprentissage. Les instructions qu'il offre en ligne peuvent être fournies **n'importe quand et n'importe où** par une gamme très vaste de solutions d'apprentissages électroniques telles que les groupes de discussions, les cours virtuels « en live », vidéo et audio, Web chat, simulations, ...etc.

En d'autre côté, la **e-formation** est liée à Internet. D'ailleurs, les anglo-saxons utilisent la notion de « web-based training ». Brandon Hall, spécialiste américain de l'e-formation, définit le e-formation comme « *un programme de formation qui est accessible via un navigateur Internet, à travers Internet ou un intranet. En utilisant le Web, ou le Web d'un intranet pour la formation, on fait référence par définition à l'environnement visuel et interactif propre à Internet.* »

La notion d'e-formation s'insère dans le cadre des formations dites « à distance ».

L'objectif étant le même : réduire le temps de présentiel, c'est-à-dire le moment où l'on regroupe les personnes à former avec le formateur dans un même lieu et à la même heure. De plus, notons que plusieurs autres modes de formation existent.

### **I.2.1.2. Autres modes de formations**

#### **La formation ouverte :**

La formation ouverte ou Open Learning, permet aux apprenants des entrées et des sorties permanentes. Elle correspond à un mode d'organisation pédagogique diversifié qui s'appuie sur des apprentissages à distance, en auto-formation. En formation ouverte, l'apprenant peut alterner des séquences individuelles et collectives. L'auto-formation est un processus par lequel l'individu détermine son itinéraire d'apprentissage (rythme, contenu, temps de travail) de façon autonome et en étant éventuellement en relation avec un tuteur ou un groupe structuré.

#### **L'enseignement à distance ou formation à distance (EAD/FAD) :**

La formation à distance est un système de formation qui permet de se former sans se déplacer sur le lieu de formation et sans la présence physique d'un formateur. La transmission des connaissances et les activités d'apprentissage se situent en dehors de la relation directe en face à face, dite « en présentiel » entre l'enseignant et l'apprenant.

La formation à distance recouvre plusieurs modalités : **cours par correspondance et e-learning.**

La formation à distance est incluse dans le concept plus général de Formation Ouverte et à Distance (FOAD).

#### **La formation ouverte et à distance (FOAD) :**

Une personne qui s'inscrit à un dispositif de formation ouverte et à distance peut, à partir de sa maison ou de son poste de travail, avec une connexion Internet, utiliser un navigateur pour accéder à la plate-forme du cours. Une fois enregistrée (identifiant et mot de passe) elle peut suivre la formation en consultant les documents pédagogiques, participer aux séances de cours en mode synchrone (chat) ou en asynchrone (forums), envoyer des questions au tuteur du module. En issus, une formation ouverte et à distance :

- se positionne avec les TIC.

*Reference ?*

- est un dispositif organisé, finalisé, reconnu par les différents acteurs.
- prend en compte la singularité des personnes dans leurs dimensions individuelle et collective.
- repose sur des situations d'apprentissage complémentaires et plurielles en termes de temps, de lieux, de médiations pédagogiques humaines et technologiques et de ressources.

En outre, un dispositif de FOAD s'appuie sur une ou plusieurs situations telles que : les cours par correspondance ; les systèmes de formation en ligne ; les centres de ressources ; les cours télédiffusés par radio ou télévision ; les campus virtuels ou classes virtuelles. Nous soulignons que, pour notre part, nous nous intéressons à la formation à distance à travers les TICE, donc médiatisée par les réseaux informatiques et entièrement à distance.

### **La Blended formation :**

Blended formation ou formation mixte, correspond à un système de formation hybride qui combine des modalités pédagogiques diversifiées, alternant formation à distance et formation en présentiel [Kassem et al, 2004].

### **I.2.2. Les enjeux du e-learning**

Le e-learning s'est beaucoup développé ces dernières années, et fait l'objet de plusieurs enjeux. Dans cette section, nous ne nous intéressons pas aux enjeux financiers et économiques. Retenons que le e-learning implique l'autonomie de l'apprenant, une certaine liberté dans le rythme de progression et d'acquisition des savoirs et des savoir faire, un changement de rôle du formateur car la connaissance est disponible en dehors de lui, il devient un médiateur. Oubahssi a précisé dans [Oubahssi, 2005] les enjeux du e-learning, nous retenons les enjeux liés : 1) à l'efficacité et à l'adaptabilité des processus d'apprentissage ; 2) à l'accès à la connaissance ; 3) à l'autonomie de l'apprenant ; 4) à l'accompagnement de l'apprenant ; 5) aux nouveaux rôles de l'enseignant ; 6) et au développement des technologies éducatives.

### **I.3. Modélisation des apprenants**

La problématique de modélisation de l'apprenant a attirée l'attention de plusieurs communautés de recherche, tel que les chercheurs travaillant dans les domaines des SIT (Systèmes Intelligents de tutorat), des SHA (Systèmes Hypermédias Adaptatifs), d'IHM



(Interaction Homme Machine), d'IA (Intelligence Artificielle) et des SR (Système de Recommandation). Cette problématique est considérée parmi les domaines d'intérêts des spécialistes de systèmes basés sur le Web. Un modèle de données décrit les caractéristiques essentielles d'un apprenant pour les usages généraux qui consistent à enregistrer et gérer les historiques, les buts et les accomplissements liés à l'apprentissage ; à décider d'engager un apprenant dans une expérience d'apprentissage ; et à découvrir des opportunités d'apprentissage pour des apprenants [IMS LIP, 2001].

### **I.3.1. Acquisition des données du modèle d'apprenant**

On distingue deux sortes globales d'acquisition et de modélisation de données liées aux apprenants :

- **Acquisition explicite** : obtenu en interrogeant directement l'apprenant par un rédacteur du profil d'apprenant ou en basant sur une série de questions conçues sous forme de questionnaire pour acquérir exactement des caractéristiques des apprenants;
- **Acquisition implicite** : par l'observation et l'inférence basées sur l'interaction d'apprenant avec le système.

#### **I.3.1.1. Modèle explicite d'apprenant**

La manière la plus simple d'acquérir le modèle d'apprenant est son interrogation directe.

C'est une façon efficace et sûre pour obtenir les informations générales de l'apprenant par des processus de diagnostic. Elle est souvent utilisée pour déclencher d'autres types d'acquisition, par exemple celles qui sont basées sur stéréotype.

Pour obtenir les informations significatives, il est nécessaire d'élaborer les questions qui assurent un diagnostic exhaustif qui pourrait exiger l'expertise multi disciplinaire et il n'est pas toujours facile de l'élaborer. Sachant que quelques apprenants ne pourraient pas aimer l'idée de diagnostic, donc il n'y a aucune garantie à laquelle les questions posées seront répondu sincèrement ou même que les questions posées sont les meilleures pour obtenir l'information désirée. Par exemple, le processus de diagnostiquer un apprenant peut être simplifié à une classification de ses résultats produits en une activité. Les résultats de l'apprenant peuvent être corrects, incorrects ou des disparus. Cependant, le diagnostic de l'apprenant s'est associé avec les problèmes de modélisation de l'apprenant, le contrôle de l'interaction tenant compte des caractéristiques dynamiques de l'apprenant, se sont des problèmes principaux d'obtention des systèmes d'éducation informatisés.

### **I.3.1.2. Modèle implicite d'apprenant**

Généralement le processus de modélisation d'apprenant implique un processus plus ou moins transparent de la perspective de l'apprenant. Afin de modéliser l'apprenant implicitement, habituellement le système collecte les données d'utilisation, se sont les données résultées de l'interaction de l'apprenant avec le système. Ainsi, la tâche la plus difficile dans la modélisation implicite est le processus d'interpréter les informations recueillies afin de produire des hypothèses valides des apprenants principalement sur leurs interactions observables avec le système.

En basant sur la durée de l'interaction avec le système, un modèle inféré d'apprenant peut être classifié comme :

- "un modèle peu profond", basé sur l'interaction relativement à court terme avec un système qui contrôle le comportement de l'apprenant, (il ne tient pas en compte des interactions de l'apprenant avec le système en sessions précédentes) ;
- "un modèle profond", qui surveille des comportements des apprenants et capture des caractéristiques de ces utilisateurs en basant sur l'interaction à long terme avec le système (il tient en compte l'histoire entière des interactions de l'apprenant avec le système).

En processus de modélisation implicite des données de l'apprenant on peut distinguer :

- les techniques de modélisation en ligne : est le processus dynamique d'extraire les caractéristiques des apprenants en basant sur les interactions de l'apprenant avec le système en temps réel;
- les techniques de modélisation off-line : est le processus en différé d'extraire des données reliées aux apprenants en basant sur « l'histoire » de l'interaction d'apprenant avec le système.

En effet, aucune des techniques précédemment citées n'est meilleure que l'autre, les systèmes impliquent des techniques en ligne ou off line ou les deux à la fois. Dans plusieurs systèmes de modélisation d'apprenant les règles d'acquisition qui déterminent ce qui peut être supposé à l'apprenant en basant soit sur son interaction avec le système soit sur ses hypothèses courantes soit sur ses données statiques soit sur le tous à la fois.

### **I.3.2. Dimensions des informations des apprenants**

Parmi un éventail de données liées aux apprenants qui peuvent être stockées dans un modèle d'apprenant. On constate neuf dimensions d'informations pour modéliser l'apprenant dans le contexte du e-learning [Baajour, 2006] :



- 1. Informations personnelles** : nom, genre, âge, culture,...etc. Certaines de ces caractéristiques affectent la manière de la présentation d'interface. Par exemple, les préférences des males et des femelles diffèrent remarquablement en termes d'attitudes, de media et de stratégies de recherche d'information.
- 2. Préférences de traitement de l'information** : elles se réfèrent à des habitudes de traitement de l'information et elles ont un impact sur les compétences et habilités des apprenants, telles que les manières préférées de percevoir l'information et de résolution des problèmes. Elles peuvent être utilisées pour personnaliser le support de navigation, la structure et la présentation du contenu et des résultats de recherche.
- 3. Spécifications du matériel** : elles concernent le matériel utilisé pour accéder à un espace d'informations et effectuer des services de personnalisation en terme de présentation d'écran et de limitations de la largeur de bande.
- 4. Contexte physique** : cette dimension capture l'environnement physique d'où l'apprenant accède au système (bureau, maison,...etc.), elle peut être utilisée pour déduire les buts de cet apprenant et adapter le contenu par conséquent.
- 5. Historique de l'utilisateur (trace)** : cette dimension capture les interactions passées de l'apprenant avec le système et peut être utilisées pour personnaliser n'importe quelle catégorie de services sous hypothèse qu'un apprenant va se comporter dans un avenir immédiat de la même manière que s'est comporté dans le passé immédiat.  
Entre d'autres données, elle peut inclure des habitudes de navigation ou des pages visitées qui contiennent des indicateurs aux mots-clés spécifiques.
- 6. Intérêts et préférences de contenu** : elles sont habituellement fournies sous forme de mots-clés ou sujets d'intérêt de l'utilisateur et elles peuvent être utilisées pour filtrer le contenu.
- 7. Motivation** : cette dimension indique pourquoi l'apprenant cherche des informations dans une session particulière. Par exemple, l'activité de recherche des informations liées à un cours pour un apprenant n'est pas pareille à celle de préparation d'un rapport pour un administratif.
- 8. Expérience d'utilisation du système** : elle montre les connaissances d'apprenant préalables qui sont liées à un espace d'informations (ex : niveau de qualifications d'ordinateur, expérience avec d'autres systèmes basés sur le Web). Cette information peut être utilisée pour personnaliser la navigation, les résultats de recherche ou offrir un help intelligent. Par exemple, l'expérience d'utilisation du système doit être liée avec les familiarités des apprenants avec les présentations et les fonctionnalités du portail d'une bibliothèque, ou avec leurs familiarités avec certaines fonctionnalités d'un portail éducatif.

**9. Connaissances de base (prérequis) :** cette dimension a une relation avec le niveau existant de compréhension d'un apprenant des connaissances d'un domaine particulier. Sachant que le niveau d'expertise d'apprenant peut varier avec le domaine ce qui influence sur le comportement de navigation, il mène aux problèmes de désorientation [Ouraiba, 2008].

En effet, l'objectif de développer un modèle d'apprenant est d'offrir des systèmes adaptatifs avec des informations nécessaires à la personnalisation du contenu, de structure et de présentation de la matière pédagogique selon notamment les compétences, les objectifs et les préférences des apprenants. Ces informations doivent être communes dans tous systèmes sachant que la plupart des standards des schémas du modèle d'apprenant ont été développés pour permettre l'interopérabilité entre les systèmes et pour réduire le plus possible la duplication de données dans plusieurs profils d'apprenants.

Néanmoins, les schémas des standards sont concentrés beaucoup plus sur le stockage et le transfert des données pour aider les administrations des institutions éducatives (par exemple en cas de déplacement des apprenants entre les institutions) et non plus de gérer les données qui peuvent être utilisées pour fournir les accès personnalisés au contenu d'apprentissage. Ceci signifie que certaines données nécessaires à la personnalisation ne sont pas incluses dans les standards proposés.

Or, jusqu'à maintenant aucun des schémas proposés ne semble avoir vraiment décollé comme la norme d'industrie, probablement parce que chacune des spécifications a ses inconvénients et aucune n'est généralement suffisante toute seule.

Maintenant nous passons en revue certains standards et spécifications courantes pour l'enregistrement et le stockage des informations d'apprenant. La liste n'est pas exhaustive, mais elle représente la gamme des fameux modèles.

### **I.3.3. Présentation des principaux modèles existants**

#### **I.3.3.1. ANSI TS 130 Student Educational Record**

La norme ANSI TS130 contient le format et établit le contenu de données d'un ensemble de transaction du registre d'éducation d'étudiant (transcription) pour l'usage dans le contexte de l'environnement de l'échange de données électroniques (EDI). La transcription d'étudiant est employée par des écoles et par les établissements éducatifs post-secondaires pour transmettre les enregistrements et les historiques courants de l'accomplissement éducatif et de toute autre information significative liée aux étudiants inscrits aux écoles et aux établissements d'enseignement. La transcription peut être envoyée à d'autres établissements éducatifs, à d'autres agences, ou



aux employeurs. La transcription d'étudiant contient l'histoire personnelle et les informations d'identification de l'étudiant, son statut scolaire courant, des cours accomplis, des résultats d'évaluation, les informations de santé ...etc. [IMS LIP, 2001] Mais, cette spécification reste uniquement comme un outil administratif.

### **I.3.3.2. vCard**

Le consortium de courrier d'Internet maintient le format de «vCard (<http://www.imc.org/>)». Le modèle vCard couvre l'essentiel des informations personnelles et des affaires en jugeant les informations habituellement trouvées sur une carte de visite professionnelle. Ce modèle représente une spécification basée sur des standards et peut être (et est) utilisée en tant qu'une base pour des profils plus impliqués d'apprenant ou d'utilisateur d'une manière générale. De plus, cette spécification convienne comme un format d'échange ouvert d'information personnelle entre les systèmes. Mais, elle manque de n'importe quelle information utile pour la personnalisation [IMS LIP, 2001].

### **I.4.3.3. eduPerson**

Les universités des USA emploient le modèle d'eduPerson pour faciliter la communication et transférer des informations liées aux personnes impliquées entre établissement d'enseignement supérieur (personnel et étudiant). Cette spécification est une classe d'objet de LDAP (Lightweight Directory Access Protocol) qui inclut les attributs couramment utilisés d'une personne dans l'enseignement supérieur. Elle comporte quelques attributs additionnels sur ceux inclus dans vCard (tel que l'affiliation, le droit, la langue préférée), mais également elle reste uniquement comme un outil administratif [UCAID, 2002].

### **I.3.3.4. Educational Modelling Languages (EMLs)**

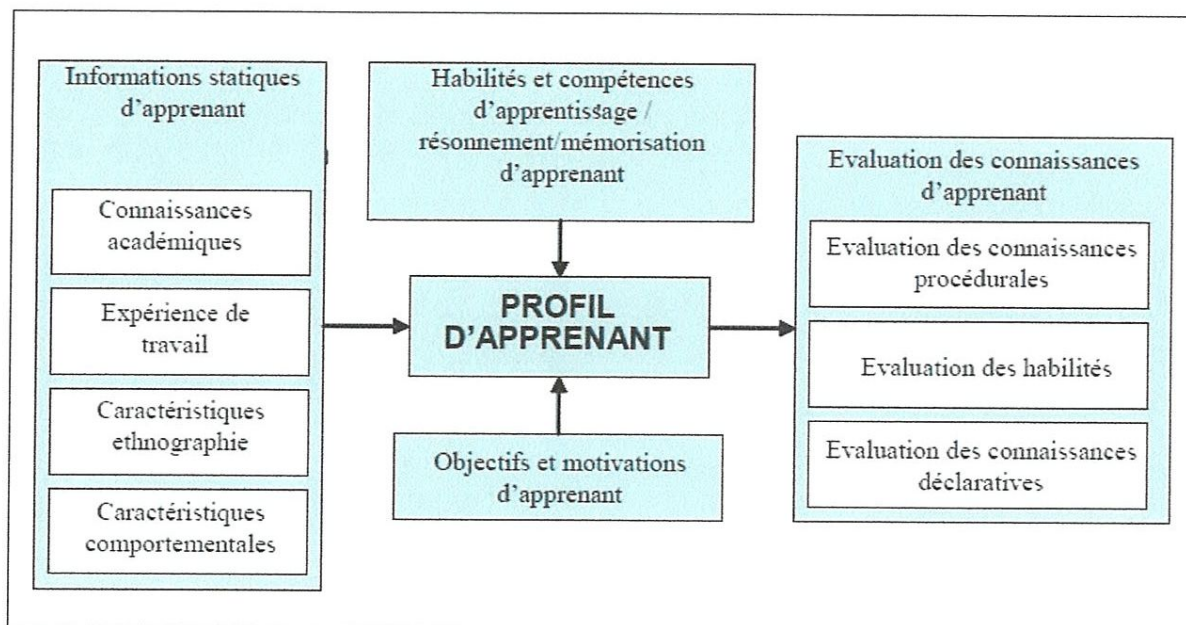
L'aperçu de EML du système de standardisation CEN/ISSS (*Comité européen de normalisation/Information Society Standardization System*) de la société d'information de CEN définit un EML comme : « *un modèle d'information sémantique ; et l'attachement et la description du contenu et du processus dans une 'unité d'apprentissage' d'après une perspective pédagogique afin de permettre la réutilisation et l'interopérabilité* ». Six langages sont passés en revue, tous ont des attaches de XML (*Extensible Markup Language*), et (3 sur 6) ont également des attaches de SGML (*Standard Generalized Markup Language*). Toutes les propositions se concentrent principalement sur la description des ressources

d'apprentissage plutôt que sur les personnes impliquées dans l'apprentissage, bien qu'un EML devrait permettre la modélisation des deux.

EML d'OUNL (Open University of the Netherlands) et aussi de PALO modélise des personnes en termes de rôles qu'ils jouent (à quelles activités ils contribuent) pour modéliser le déroulement des opérations. Ce genre de modélisation d'apprenant peut être certainement utilisé en créant des parcours personnalisés, malgré que des informations supplémentaires telles que des buts courants, des expériences et des préférences d'apprenant (qui ne sont pas couvertes par la modélisation de déroulement des opérations) sont également souhaitables dans un modèle d'apprenant [Rodríguez, 2002].

### I.3.3.5. Le modèle d'apprenant de Chen et Mizoguchi

Ce modèle est présenté sous forme de quatre relations entre : les informations statiques ; les objectifs et les motivations ; les habilités et les compétences ; et l'évaluation des connaissances de l'apprenant. La figure 1.1 ci-dessous montre d'une manière détaillée ce modèle :



**Figure 1.1.** Le modèle d'apprenant de Chen and Mizoguchi [Chen et Mizoguchi, 1999].

Ce modèle est exploité dans les travaux liés à la communication des agents dans le domaine des systèmes éducatifs intelligents (Intelligent Educational Systems IES) par Chen et Mizoguchi [Chen et Mizoguchi, 1999]. L'objectif principal de ce modèle est de définir les entités, les relations et les fonctions nécessaires à la description des procédures de la

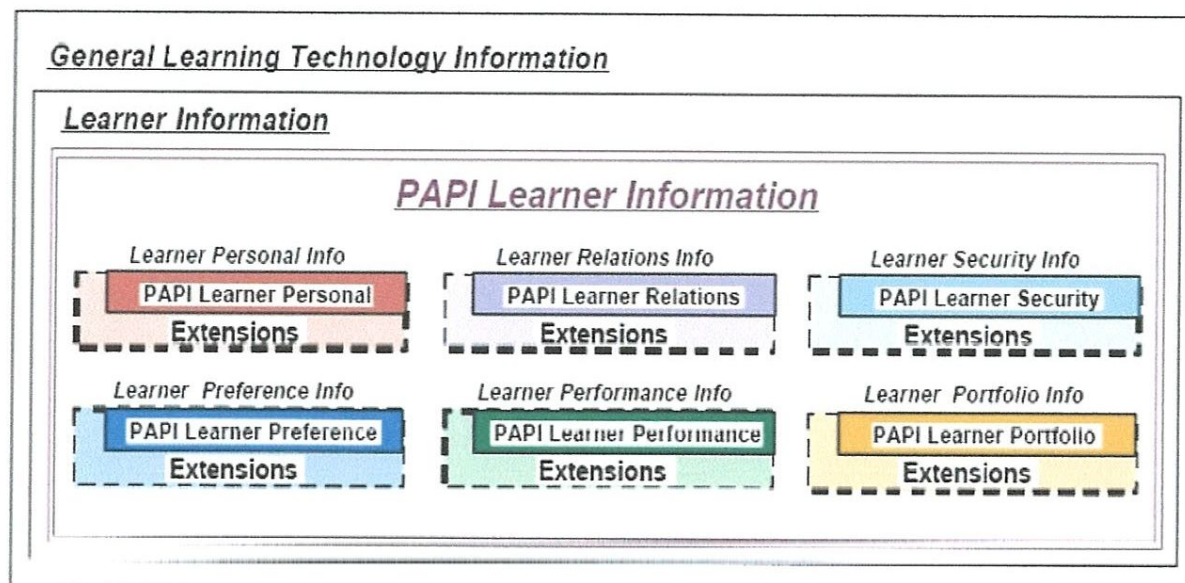


modélisation d'apprenant, de la prise des décisions éducatives, de l'interaction avec l'apprenant et de la communication entre les agents.

Cependant, comme il est mentionné dans (la figure 1.1), ce modèle comporte les informations détaillées sur la notion de « l'évaluation de l'apprenant ». Cet aspect est lié à des caractéristiques cognitives d'apprenant lorsqu'il est exposé à une pièce de connaissance telle que la mémorisation d'apprenant et sa capacité d'accélération d'apprentissage et de raisonnement qui sont décrites exhaustivement. Ainsi, parmi les avantages de ce modèle est qu'il a été déjà validé et utilisé. De plus, il est techniquement orienté vers son intégration avec les agents et les outils logiciels. D'un autre côté, parmi les inconvénients est qu'il ne prend pas en considération en issues les objets d'apprentissage et les interactions de l'apprenant avec son environnement et avec d'autres personnes.

### 1.3.3.6. IEEE Public and Private Information – Learner specification(PAPI)

PAPI (*Public and Private Information for Learners*) est le standard proposé par IEEE qui est un sous ensemble des informations générales des technologies de formation [PAPI, 2000], il décrit un sous ensemble des informations d'apprenant. Cependant, le terme « profile d'apprenant » est le nom générique, alors que PAPI est seulement une description spécifique de ce profile d'apprenant. Ainsi, il spécifie simultanément la syntaxe et la sémantique de modèle d'apprenant qui peut être utilisé pour caractériser l'apprenant. La figure 1.2 ci-dessous présente ce modèle :



*Figure 1.2. Le modèle d'apprenant PAPI [PAPI, 2000].*

PAPI identifie six types d'informations de profile d'apprenant qui sont décrits brièvement comme suite :

- 1. Les informations personnelles de l'apprenant :** sont principalement destinées à des usages d'administration et ne sont pas liées directement à la capture et au enregistrement des performances de l'apprenant. Pour cela ce type d'informations est privé et sécurisé. Par exemple : nom, adresse, numéro de sécurité sociale...etc.
- 2. Les informations des relations de l'apprenant :** concernent les relations de l'apprenant avec d'autre utilisateur du système de technologie de formation tels que l'enseignant, le tuteur, et d'autres apprenants.
- 3. Les informations de sécurité de l'apprenant :** liées à des références de sécurité d'apprenant tels que : mot de passe, challenge/réponses, clés privées, clés publiques et biométriques...etc.
- 4. Les informations des préférences de l'apprenant :** décrivent les préférences de l'apprenant afin d'améliorer l'interaction homme/machine tels que : appareil d'E/S utile ou inutile, styles d'apprentissage, limitations physiques...etc.
- 5. Les informations de performance de l'apprenant :** liées au historique de l'apprenant ; à son travail courant ; ou à ses objectifs futurs tels que : des rapports, des fichiers log, des gades,...etc. Ces informations sont créées et utilisées par des composants de la technologie de formation pour fournir des expériences d'apprentissage perfectionnées et optimisées.
- 6. Les informations de portfolio de l'apprenant :** représentent une collection représentative des travaux de l'apprenant ou références à eux. Ces informations sont exploitées pour l'illustration et la justification des habilités et des accomplissements d'apprenant.

Ainsi, ce modèle est proposé en répondant à l'exigence de définir un standard de modélisation de profile d'apprenant dans le domaine des technologies de l'information d'apprentissage. Bien que, les données de ce modèle peuvent être étendues par des informations additionnelles afin d'avoir un modèle conformant et plus représentatif.

En effet, ce modèle ne décrit pas toutes les informations possibles de l'apprenant, mais il inclut uniquement le minimum nécessaire des informations pour satisfaire des besoins opérationnels et afin d'être plus portable et pour permettre l'extension de ces informations. Ceci explique l'objet d'une évolution faite par IMS dans son nouveau standard IMS LIP [IMS LIP, 2001].

#### **I.3.3.7. IMS Learner Information Package (IMS LIP)**

La spécification « *Instructional Management System Learner Information Package* » ou IMS LIP représente une collection d'informations liées à l'apprenant (individuel ou groupe) ou au producteur du contenu d'apprentissage (créateur, fournisseur ou vendeur). Cette spécification



permette l'interopérabilité des systèmes d'information d'apprenant qui sont basés sur Internet avec d'autres systèmes ou environnements Internet supportant d'apprentissage. Autrement, elle définit une structure XML pour l'échange des données apprenant entre systèmes coopérants tels que : Les systèmes de gestion d'apprentissage, les systèmes des ressources humaines, les systèmes de gestion des données des apprenants, les systèmes de gestion des connaissances, et d'autres systèmes utilisant les processus d'apprentissage.

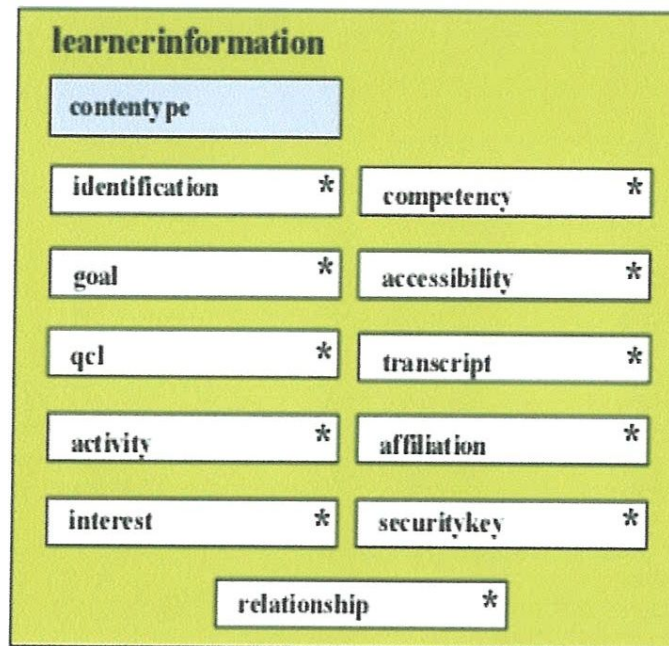
Cette spécification est proposée pour satisfaire les exigences suivantes [IMS LIP, 2001]:

- Des informations distribuées : plusieurs systèmes qui stockent et partagent les informations des apprenants ;
- Scalabilité : afin de supporter des systèmes à grande échelle et maîtriser ses granularités ;
- Privatisation et protection de données : pour assurer l'intégrité de données ;
- Flexibilité et référencement externe : pour comprendre et intégrer les données de sources différentes tels que : les objectifs et les historiques d'apprentissage.

Nous notons que l'élaboration du modèle IMS LIP s'est beaucoup basée sur la spécification de PAPI. La première version (1.0) des spécifications de IMS LIP a été élaborée en 2001, elle comporte trois documents :

- IMS Learner Information Package Information Model ;
- IMS Learner Information Package XML Binding;
- IMS Learner Information Packaging Best Practice & Implementation Guide.

Le paquetage LIP peut comprendre beaucoup d'éléments mais la plupart de ceux-ci sont optionnels et leur implémentation est laissée au choix de l'utilisateur. Dans cette version, le modèle IMS LIP définit une structure de données utilisateur en 11 catégories pour pouvoir être importée ou exportée entre des systèmes interopérables. La figure 1.3 ci-dessous donne une vue sur les onze catégories qui constituent ce modèle, sachant que l'élément «**Contenttype**» concerne les métas données.



*Figure 1.3. Représentation de la proposition de standard IMS LIP [IMS LIP, 2001].*

Les onze catégories du modèle IMS LIP sont représenté ci-dessous :

- 1. Identification :** Cette catégorie contient des éléments Biographiques et démographiques qui aident à identifier l'apprenant tels que son nom, son genre, son adresse, son email, ...etc.
- 2. Accessibility :** Cette catégorie permet de décrire l'accessibilité de l'apprenant tels que : ses langues, ses éventuels handicaps, ses éligibilités, ses préférence d'apprentissage en incluant des préférences cognitives (comme : style d'apprentissage), ses préférences physiques (comme : format imprimé ou non) et ses préférences technologiques (comme : plateforme particulière)...etc.
- 3. QCL :** Cette catégorie permet de décrire les qualifications, les certifications et les diplômes attribués à un apprenant.
- 4. Activity :** Cette catégorie regroupe les données sur les activités formelles ou informelles liées au travail et à la formation d'un apprenant.
- 5. Goal :** Cette catégorie contient les données sur les objectifs et les inspirations liés à l'apprentissage des apprenants.
- 6. Competency :** Cette catégorie décrit les compétences, les connaissances, les habilités associées avec la formation formelle ou informelle d'un apprenant et son expérience de travail.
- 7. Interest :** Cette catégorie regroupe les données sur les hobbies et les activités créatrices d'un apprenant.



**8. Transcript** : Cette catégorie permet de décrire les données sur les achèvements académiques de l'apprenant. La structure de cet enregistrement peut prendre plusieurs formes.

**9. Affiliation** : Cette catégorie inclut les informations sur la description de l'organisation associée à l'apprenant.

**10. Securitykey** : Cette catégorie regroupe les données de sécurité d'un apprenant, tels que les mots de passe et les codes de sécurité qui doivent être utilisés durant ses communications. Pour chaque clé de sécurité une structure différente sera utilisée.

**11. Relationship** : Cette catégorie permet de décrire les relations entre les structures de données utilisées pour stocker les données de l'apprenant employée dans ce modèle.

### **I.3.3.8. EuropeanDiplomaSupplement (EDS)**

Le modèle EDS (European Diploma Supplement) est un profil de IMS LIP. C'est le résultat des travaux du groupe de travail « Supplément au diplôme ». Son objectif est d'améliorer la «transparence» internationale et de faciliter la reconnaissance académique et professionnelle des qualifications (diplômes, acquis universitaires, certificats,...etc.). Il est destiné à décrire la nature, le niveau, le contexte, le contenu et le statut des études accomplies avec succès par l'apprenant désignée par la qualification originale à laquelle ce supplément est annexé. Récemment le profil EDS est disponible dans les onze langues officielles de l'union européenne.

Cependant, toutes les informations requises dans ce modèle sont obligatoires et sont définis à partir des adaptations sur une partie des éléments de IMS LIP. Nous notons aussi que l'analyse de ce modèle nous a donné une vue sur l'exemple d'une adaptation du IMS LIP pour un contexte qui n'est pas nécessairement lié à e-learning mais qui est lié à des formations académiques.

### **I.3.3.9. Dolog Learner Profile (Dolog LP)**

Dolog et al proposent un modèle d'apprenant en utilisant RDF (Resource Description Framework) et une ontologie d'apprenant pour avoir des systèmes d'apprentissage enrichies sémantiquement afin de fournir des services personnalisés. Leur profil d'apprenant comporte l'avantage de la flexibilité de RDF en construction du modèle d'utilisateur pour inclure les attributs des standards multiples (schémas) et de l'habilité d'ajouter autres attributs si nécessaire au modèle d'apprenant selon leur usage. L'utilisation de RDF facilite aussi l'échange de profil d'apprenant entre des systèmes dans les environnements ouverts et assurer l'interopérabilité entre eux. Dolog et al ne donnent pas un intitulé officiel à leur modèle

d'apprenant, mais dans la plupart des travaux est appelé par Dolog LP (Dolog Learner Profile). Etant donné que le but du Dolog LP est de fournir des services personnalisés. Ce modèle est basé sur la combinaison entre PAPI et IMS LIP, tel que les caractéristiques d'apprenant modélisées sont choisies d'après les deux spécifications afin de sélectionner les avantages de chacune et éviter les limitations des deux dans le contexte de personnalisation. En conséquence, ce modèle présente les informations de l'apprenant dans cinq catégories [Dolog, 2004]:

- 1. Identification** : comporte des informations liés au apprenant tels que : son nom, son téléphone, son adresse, son e-mail...etc.
- 2. Autres caractéristiques de l'apprenant** : décrivent les préférences de l'apprenant (langues, compétence, éligibilité, handicap, et n'importe quelle autre préférence), les objectif (Goal) et les hobbies (Interest) de l'apprenant sont utilisés d'après IMS LIP ;
- 3. Performance d'apprentissage** : décrit la performance de l'apprenant, son portfolio et sa certification ;
- 4. Planification des ressources humaines** : comporte des informations de département ou d'organisation dans laquelle l'apprenant est un membre;
- 5. Calendrier** : contient des informations liées aux rendez-vous, rencontres et événements aux quelles l'apprenant assiste. Cette catégorie n'est pas héritée à partir de PAPI ou IMS LIP mais plutôt est introduite dans des informations additionnelles du modèle.

En effet, on peut considère le modèle Dolog LP comme une extension d'IMS LIP, sachant qu'il ne couvre pas la caractéristique de relations, mais il est mieux utilisé dans les systèmes adaptatifs du e-learning.

#### **1.3.3.10. Friend of a Friend (FOAF)**

Le modèle FOAF est très exploité pour modéliser les apprenants dans les réseaux sociaux dans le cadre des systèmes d'apprentissage à distance et des organisations à rand échelle et multi culturelles tels que les universités virtuelles où l'apprenant est physiquement situé dans différentes localisations et leurs vies sociales sont complètement séparées par rapport à la vie académique d'université. D'après ce point de vue, le besoin de groupements sociaux est de plus en plus augmenté, tel que chaque apprenant a besoin de camarades qui partagent les mêmes intérêts, préférences, expériences d'apprentissage ou autres caractéristiques. Puisque le nombre d'apprenants dans ce type de systèmes est très grand, le groupement social doit être généré automatiquement et dynamiquement. Dans ce point là, on note l'exigence de modèle d'apprenant pour engendrer des communautés de pratique. En issue, on mentionne FOAF



comme un profile possible d'apprenant parmi les autres standards. FOAF est vocabulaire de RDF qui fournit un ensemble de propriétés et de classes pour décrire les personnes, les documents et les organisations .Il est développé pour construire des communautés et des groupements sociaux, donc avoir des relations pour exprimer les membres des communautés [Dumbill, 2002].

Le modèle de données FOAF distingue cinq catégories de description d'apprenant :

**1. Basiques de FOAF** : comportent des informations basiques sur l'apprenant telque : son nom, son e-mail, ses images, sa page personnelle,...etc.

**2. Informations personnelles** : incluent plus d'informations personnelle tels que :

Weblog, intérêts, projets, home page scolaire, publications, ...etc.

**3. Comptes en ligne** : contiennent des informations liées aux comptes del'apprenant comme : compte en ligne, intitulé du compte, compte du chat,compte de forum,...etc.

**4. Projets et groupes** : incluent des informations liées aux projets, groupes ou organisations dans lesquelles l'apprenant est un membre tels que : projet, groupe, membre, financé par, thème...etc.

**5. Documents et images** : comportent des informations liées à un document ou à une image tels que : document profile personnel, logo, image,...etc.

En effet, on constate que FOAF décrit les relations de l'apprenant avec les autres enfocalisant notamment sur son profil cognitif, mais pratiquement il ne couvre pas tous les achèvement et les objectifs de l'apprenant.

#### **I.4. Modélisation des ressources pédagogiques**

Le développement de la formation en ligne et l'utilisation du Web pour la recherche et la diffusion des ressources de formation rendent indispensables le partage et l'échange de ressources d'enseignement entre différents viviers de ressources via des catalogues.

Partager et échanger des ressources nécessite le respect de mécanismes uniques par les différents systèmes. Cette unicité est assurée par le respect des normes et des standards.

Nous présentons les plus fameux dans le suivant.

## I.4.1. Présentation des principaux modèles existants

### I.4.1.1. Le modèle DublinCore

En mars 1995, et suite à un séminaire organisé à Dublin (Ohio), les premières bases d'un système de descripteurs de documents ont été proposées, comprenant 15 catégories générales (titre, créateur, sujet, description, éditeur, contributeur, date, type, format, identifiant, source, langage, relation, étendue, droits). Ces catégories forment le noyau de nombreux systèmes de classification. C'est pourquoi ils sont appelés « Dublin Core Metadata ». A la suite de ce séminaire, est créée la Dublin Core Metadata Initiative, pour développer ces spécifications.

#### I.4.1.1.1. Présentation des catégories du DublinCore

Les 15 éléments de ce modèle sont [Dcml, 2004] :

- 1. Titre** : Nom propre donné, par l'auteur, à la ressource décrite.
- 2. Créateur** : La personne physique ou collectivité identifiable créatrice de la ressource, et responsable du contenu intellectuel et/ou artistique de celle-ci.
- 3. Sujet** : Le sujet du contenu de la ressource. Autrement dit : Notions et concepts sur lesquels le document apporte une information significative, exprimés en mots-clés libres, descripteurs (vocabulaire contrôlé) et indices (classification décimale).
- 4. Description** : La description du contenu de la ressource sous forme de résumé analytique indicatif, ou de présentation, signalant le ou les thèmes du document.
- 5. Editeur** : La structure qui a rendu disponible la ressource dans sa forme actuelle.
- 6. Contributeur** : Les autres personnes ou organisations impliquées dans le processus de production de la ressource (éditeur, traducteur, etc.).
- 7. Date** : Dates associées au cycle de vie de la ressource (création, publication, modification).
- 8. Type** : La forme de la ressource déterminée en fonction de son organisation, de ses composants ou de son usage.
- 9. Format** : La forme physique de la ressource.
- 10. Identifiant** : Une référence permettant d'identifier et donc de localiser une ressource sans ambiguïté.
- 11. Source** : Les informations à propos d'un document d'origine ayant servi à la production du document présent.
- 12. Langage** : La langue du contenu de la ressource.
- 13. Relation** : Fait référence à une ressource liée à la ressource présente.



**14. Etendue :** Les caractéristiques spatiales et temporelles du document (lieu et date de l'action typiquement).

**15. Droits :** Représente l'identification des personnes morales et physiques concernées par les droits de tous types générés par la ressource. Informations sur ces droits en référence à la législation relative à la notion de propriété intellectuelle.

En effet, le DublinCore n'a pas été conçu spécifiquement pour le e-learning, plusieurs acteurs jugent qu'il n'est pas suffisamment complet et adapté au domaine de l'enseignement. Par conséquent, il a donné lieu à des adaptations pour l'application aux documents pédagogiques.

#### **I.4.1.2. Le modèle ARIADNE**

Dans sa philosophie, le système ARIADNE cherche à diminuer le travail de production du matériel pédagogique par la mise en commun des matériels existants et la réutilisation systématique de composants pédagogiques par les enseignants. La réutilisation des composants pédagogiques peut prendre deux formes distinctes : d'une part, la création de nouveaux modules constitués d'une mosaïque de fragments de modules empruntés aux quels l'auteur peut ajouter de nouveaux éléments ; et d'autre part, la réalisation d'une nouvelle présentation d'un cours existant obtenue par un réarrangement systématique de ses composants. Cela implique que tout auteur participant à l'expérience du vivier autorise l'usage et la modification des composants qu'il y introduit. En échange, il peut lui-même bénéficier de la réciproque. Cette approche favorise aussi l'harmonisation des contenus dispensés aux apprenants [Vidal et al, 2004].

##### **I.4.1.2.2. Présentation des catégories du ARIADNE**

Le modèle ARIADNE, regroupe les différentes métas-données associées à toute ressource pédagogique à indexer dans le vivier du système. Ses descripteurs sont classés dans sept catégories suivantes [Najjar et al, 2003] :

- 1. Les données générales :** Regroupe les informations générales qui décrivent la ressource pédagogique, tels que, le titre de la ressource, la langue d'utilisation de la ressource, ...etc.
- 2. Les données sémantiques :** Regroupe les informations qui décrivent la classification sémantique de la ressource, tels que le type du domaine, la discipline, ...etc.
- 3. Les données pédagogiques :** Regroupe les informations qui décrivent les caractéristiques pédagogiques des ressources tels que, la densité sémantique, le niveau d'interactivité, ...etc.

**4. Les caractéristiques techniques :** Regroupe les informations qui décrivent les caractéristiques et les exigences techniques des ressources tels que le format de la ressource, sa taille, le système d'exploitation, la taille mémoire,...etc.

**5. Les conditions d'utilisation :** Regroupe les informations sur les conditions d'utilisation de la ressource.

**6. Les métas-métadonnées :** Regroupe les données qui décrivent les informations générales sur les métadonnées tels que l'identificateur des métas-données, le créateur et la date de création de la métadonnées,...etc.

**7. Les données optionnelles (annotations) :** Regroupe les données sur les personnes ou les organisations qui ont annoté les ressources tels que l'annotateur, la langue d'annotation, la date d'annotation,...etc.

En effet, ARIADNE a beaucoup contribué aux solutions dans le domaine de la portabilité des contenus du e-learning. Ses acteurs participent activement à la résolution des problèmes d'interopérabilité entre les différents systèmes qui préoccupent tous les acteurs du e-learning. Son modèle de métadonnées, représente l'un des principaux éléments de base sur laquelle est fondé le standard LOM que nous allons présenter dans le paragraphe qui suit.

#### **I.4.1.3. Le modèle LOM**

Le modèle LOM (Learning Object Metadata) est le résultat d'un projet de l'IEEE-LTSC visant à proposer un modèle standard de métadonnées permettant de décrire et référencer toute entité pédagogique numérique. Ce Projet est fondé sur les travaux préliminaires de l'IMS et d'ARIADNE. Le groupe de travail est également constitué de l'ensemble des acteurs actuels du domaine en terme de normalisation, tels que AICC, Dublin Core, SCORM,...etc.

LOM a été adopté comme standard par le LTSC de l'IEEE en 2002 sous l'identification 1484.12.1-2002, également désigné par le nom « IMS Learning Resource Meta-data ». Son objectif est de:

- faciliter la recherche, l'évaluation, l'acquisition et l'utilisation des objets pédagogiques par les apprenants, les enseignants ou les processus logiciels automatisés ;
- définir un schéma conceptuel de données qui détermine la structure d'une instance de métadonnées pour un objet pédagogique ;
- déterminer un schéma élémentaire extensible capable d'intégrer les modifications découlant de l'expérience acquise afin de faciliter, par exemple, l'ordonnement automatique et évolutif des objets pédagogiques par des agents.



Dans son modèle, LOM propose 45 éléments descriptifs de premier niveau regroupés en neuf catégories. Dans chacune d'entre elles plusieurs éléments peuvent être répétés parfois de façon récursive.

Pour donner leurs valeurs aux éléments descriptifs, le LOM utilise plusieurs ensembles de balises « outils » : langstring (balise <langstring>) permet de préciser des valeurs selon plusieurs langues (avec l'attribut xml:lang), vocabulary (balises <source>,<value>) qui permet de les choisir en accord avec un vocabulaire défini, date (balises<datetime>,<description>) et vcard (balise <vcard>) pour la « carte de visite virtuelle »[LOM, 2002].

#### **I.4.1.3.1. Présentation des catégories du LOM**

Une brève description des catégories du LOM v1.0 est comme suite :

- 1. Description général (general) :** Cette catégorie décrit l'objet pédagogique dans son ensemble. On y trouve des données sur l'identifiant de l'objet, son titre, sa description, la liste des langues utilisées, une liste de mots clés, l'étendue de la ressource (temps, géographie, culture ...), le type de structure (collection, linéaire, hiérarchique ...), son niveau de granularité (de 1 à 4, 1 désignant un cours entier).
- 2. Cycle de vie (life cycle) :** Cette catégorie permet de décrire les caractéristiques relatives à l'historique et à l'état courant de l'objet pédagogique (draft, final...), les personnes qui l'ont modifié, à quelle date ainsi que leur rôle (author, publisher, instructional designer...). Cette partie décrit la liste complète des modifications ou cycle de révision.
- 3. Méta-méta-données (meta-meta-data) :** C'est un ensemble de métadonnées sur les métadonnées décrivant l'objet pédagogique. Cet ensemble décrit le schéma ou la spécification utilisée. Il est possible de satisfaire à plusieurs schémas et de définir des liens dans un système de catalogue connu.
- 4. Les informations techniques (technical) :** Cette catégorie décrit les exigences techniques en terme de navigateur (type, version), de système d'exploitation ou les caractéristiques comme le type des données ou format (permettant d'identifier les logiciels nécessaires pour les lire), la taille de l'objet numérique (en octets), sa localisation physique (URL Uniform Resource Locator ou URI UR Identifier), des informations pour installer l'objet pédagogique et sa durée (en particulier pour les fichiers de type son, animation ou vidéo).
- 5. La partie pédagogique (educational) :** Cette catégorie permet de définir les conditions d'utilisation de la ressource : Comment celle-ci doit être utilisée, savoir quel est son type (exercice, figure, index...), son niveau d'interactivité (de très faible à très élevée), à qui s'adresse la ressource (apprenant, enseignant, auteur...), le contexte (université, formation

professionnelle, école primaire...) ou la tranche d'âge à laquelle s'adresse la ressource. C'est souvent par ces caractéristiques que l'on améliore l'exploitation du contenu pédagogique.

**6. La gestion des droits (rights) :** Cette catégorie concerne les droits (copyright) liés à la ressource pédagogique, éventuellement son coût.

**7. L'aspect relationnel (relation) :** Cette catégorie couvre les différents types de lien entre les objets qui peuvent être de granularité différente. Il est en particulier possible de décrire différents niveaux d'agrégation grâce aux relations d'appartenance ou de décomposition : L'élément « Niveau d'agrégation » propose quatre niveaux et peut être appliqué implicitement à deux types d'objet : Une ressource numérique ou un élément de structuration pédagogique.

**8. Annotation :** Cette partie regroupe les commentaires sur l'utilisation de la ressource ainsi que les détails relatifs à l'auteur et à la date de création des commentaires.

**9. Classification :** Cette catégorie indique l'appartenance de la ressource à une ou plusieurs instances de classification.

En effet, le standard LOM fournit une base concrète de départ pour la normalisation et l'indexation des ressources d'enseignement utilisées dans les systèmes de gestion de la formation. Dans l'annexe 2, nous présentons les attributs du modèle LOM en détail.

En outre, le modèle LOM a été adopté par la plupart des organismes (IEEE LTSC, IMS, ADL, ARIADNE) et le plus souvent adapté dans le cadre des profils d'application tels que : CanCore, Celebrate, Normetic, ... etc.

## **I.5. Conclusion**

Dans ce premier chapitre, nous avons précisé tous d'abord le concept du e-learning et ses enjeux. Par la suite, nous avons abordé un premier souci qui concerne des environnements d'apprentissage, il s'agit de la modélisation des apprenants. Nous avons alors commencé par la présentation des deux façons d'acquies des données des apprenants. En suite, nous avons déterminé les principales dimensions d'un modèle d'apprenant. Par conséquent, nous présentons les différents modèles d'apprenant existants. Dans la dernière section de ce chapitre nous avons entamé la modélisation des ressources pédagogiques. Nous avons alors présenté les modèles existants des métadonnées d'un objet d'apprentissage.

Le deuxième chapitre qui suit sera consacré à la présentation des systèmes de recommandation qui sont appelés aussi les systèmes de filtrage d'information.

# Chapitre 2 :

# Les systèmes de

# recommandation

---



## **Chapitre 2 : Les systèmes de recommandation**

### **II.1. Introduction**

Les contenus sur le Web sont largement volumineux et non structuré. Cela implique trop de consommation de temps et d'effort pour être pratique en recherchant l'information pertinente. Alors que, dans ce contexte les systèmes de recommandation qui sont appelés aussi les systèmes du filtrage d'information trouvent leur place. Ces systèmes recommandent des pages Web, des news, des objets d'apprentissage, ou n'importe quel autre genre du contenu aux utilisateurs suivant leurs propres profils [Middleton, 2004].

### **II.2. Définition du filtrage d'information**

Le filtrage d'information connu aussi sous le nom de la collecte passive d'information, est le processus permettant à partir d'un large volume d'informations dynamiques, d'extraire et de présenter les seuls objets intéressants un utilisateur ou un groupe d'utilisateurs ayant des centres d'intérêts relativement semblables appelés profils. Ce processus fait parvenir continuellement les informations aux personnes qui en ont besoin à partir d'un grand volume d'information dynamique, au lieu de les laisser dépenser leur temps à chercher ce que les intéressent. En outre, Favier dans a défini le filtrage de l'information comme étant un procédé visant à reproduire automatiquement l'attitude spontanément sélective d'un utilisateur face à l'offre informationnelle qui change à chaque moment.

En effet, le but du filtrage d'information est d'augmenter la quantité d'informations pertinentes collectées à partir de différentes sources en ciblant l'information vraiment appropriée suivant les besoins de connaissance établis par l'utilisateur[Favier, 2001].

### **II.3. Brève Historique des systèmes du filtrage d'information**

Des systèmes du filtrage d'information on connaît plusieurs évolutions pendant différentes périodes : Business Intelligent System (1957, Luhn) ; Diffusion Sélective d'Information 'SDI' (Luhn) ; Création de SIG-SDI (Special Interest Group on SDI de American Society of Information Science), 10 ans plus tard ; Expression FI (Denning, 1982) ; trois formes: cognitif, économique et social (Malone, 1987) ;Tâche de filtrage dans TREC (Text Retrieval

Conference) (NIST, National Institute of Standards and Technology, 1992) ; Numéro spécial ACM (9 articles, Décembre 1992)[Ouraiba, 2008].

#### **II.4. Domaines d'application du filtrage d'information**

Les domaines d'application du filtrage d'information sont divers et d'une importance capitale sur le plan économique, citons parmi eux : Mailing list, Usenet News, E-mails, services de dissémination de l'information et le filtrage des informations sur le Web...etc. Ce dernier est réalisé par la détermination des contenus à envoyer au visiteur d'un site Internet en fonction des critères spécifiques qui lui sont propres et qui constituent son profil.

Ainsi, les systèmes de recommandation ont été utilisés pour plusieurs domaines. Ils sont émergés dans les domaines du e-commerce, les bibliothèques numériques et la gestion des connaissances, ils fournissent des informations personnalisées en basant sur les préférences de l'utilisateur. Avec l'apparition du commerce électronique, les systèmes de recommandation ont beaucoup utilisé pour recommander des produits qu'un client achètera très probablement : les films, les programmes de TV ou la musique qu'un utilisateur trouvera agréable comme par exemple le système Ringo qui recommande la musique à des individus. Au de là de ce dernier, plusieurs systèmes de recommandation ont été réalisés dans différents domaines, parmi eux nous citons à titre d'exemple : Letizia qui assiste l'internaute à naviguer en proposant des liens, il est réalisé par Lieberman en 1995 ; Web Watcher dans le domaine de la catégorisation du texte ; Fab, il capte le profil d'utilisateur en lui proposer d'ordonnancer des pages ; Web Mate[Chen et Sycara, 1998];Morse par BT Telecom ; Poly Lens par O'Connor en 2001 ; News Weeder , Info Finder, Sys kill et Webert,...etc. ; et aussi les sites Web notamment du e-commerce tel que « [www.amazon.com](http://www.amazon.com) » de vente des livres en ligne.

#### **II.5. Caractéristiques d'un système du filtrage d'information**

Un système du filtrage d'information :

- exploite un grand volume de données entrant, transmis par des sources distantes ;
- se base principalement sur le profil de l'utilisateur ou d'un groupe d'utilisateurs ;
- diffuse que les informations en adéquation avec le profil l'utilisateur [Abbes, 1999].

Des systèmes du filtrage d'information se caractérisent par :



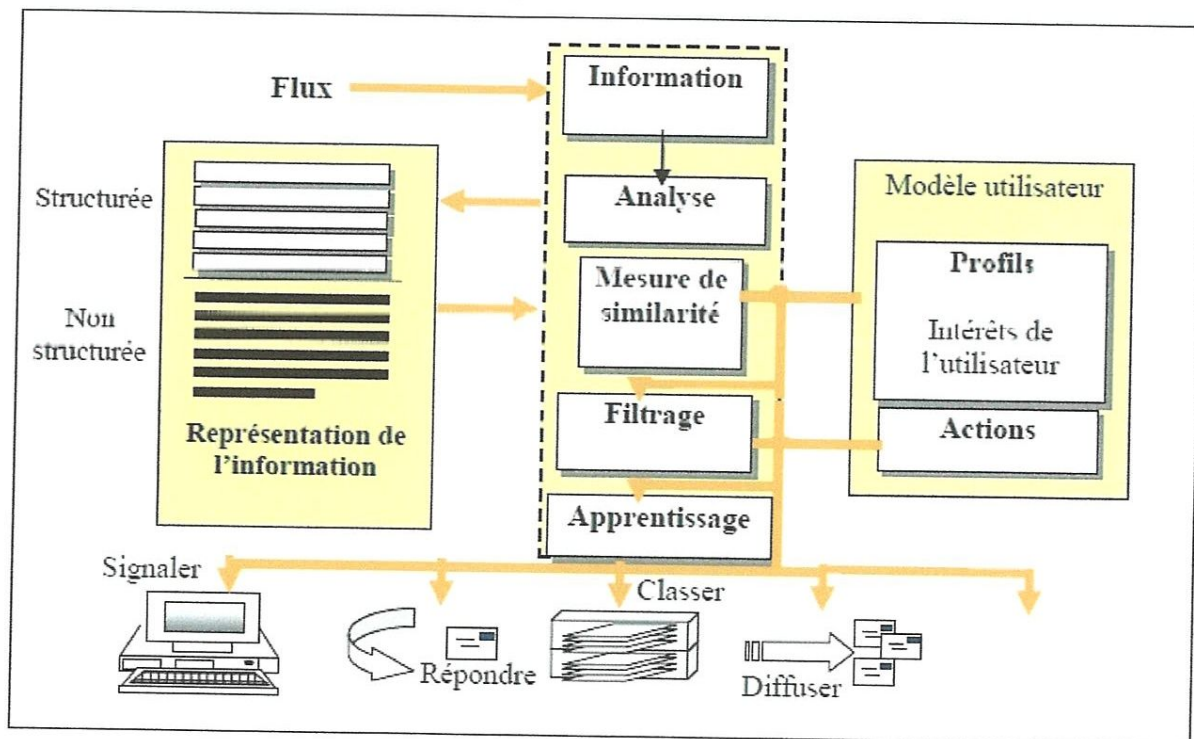
- La façon de représentation de l'information telle que la représentation vectorielle sous forme de « sac de mots » (termes sans structure) ; ou unités syntaxiques (termes + structures syntaxiques) ; ou unités sémantiques ou conceptuelles (sens des termes).
- Les techniques de représentation : DF, IG,  $\chi^2$ , MI,...etc.
- Les techniques de pondération : binaire, TF, IDF, TFIDF,...etc.
- Les techniques de réduction : ACP, LSI, classification, clustering,...etc.
- Les méthodes de similarité : S-Matching, Dice, Jaccard, cosinus,...etc.

Les méthodes d'évaluation du filtrage :

- Quantificateurs : Précision, Rappel, Fallout, Mesure\_F, Utilité, etc.
- Programmes d'évaluation : MUC, TREC, TIPSTER, TDT, AMARYLLIS, EXIBUM...

## II.6. Mécanisme du fonctionnement d'un système du filtrage d'information

Un système du filtrage d'information est défini par son modèle de représentation des profils utilisateurs, son modèle de représentation des objets et sa fonction de prédiction sur la pertinence des objets reçus. La figure 2.1 ci-dessous montre les éléments de base d'un système du filtrage d'information [Nouali, 2006] :



*Figure 2.1. Eléments de base d'un système du filtrage d'information [Nouali, 2006].*

Le filtrage achemine des objets entrant vers un utilisateur ou un groupe d'utilisateur en se basant sur leurs profils, sortes de requêtes à long terme. Il utilise plusieurs mécanismes de la

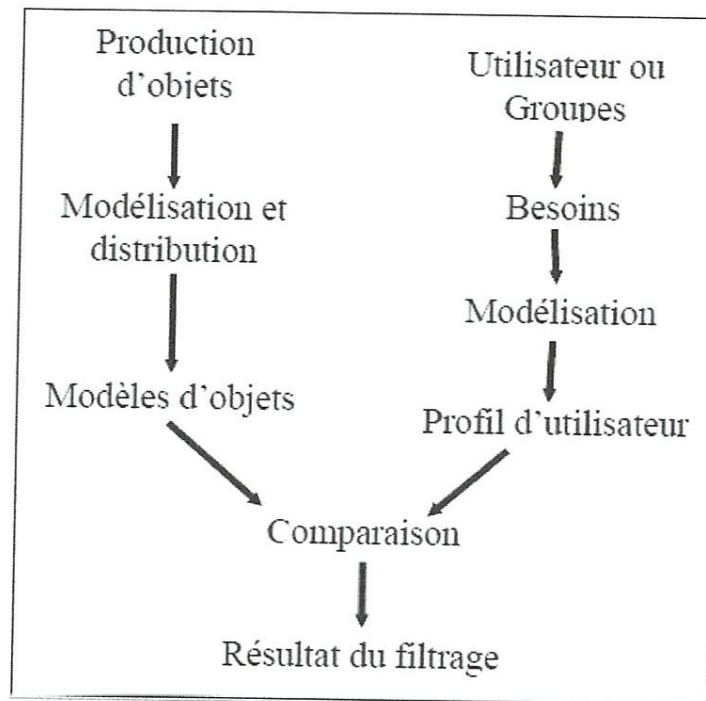


recherche d'information du fait qu'il opère une comparaison entre les métas-données des objets et le profil de l'utilisateur. Le déclenchement du processus du filtrage commence par l'expression des besoins de l'utilisateur qui vont constituer son profil initial. Ces besoins peuvent être stables, périodiques ou à long terme.

En d'autre coté les producteurs des objets diffusent leurs produits dès qu'ils sont générés, une modélisation adéquate sera alors associée à chaque objet diffusé.

L'appariement est effectué par la suite entre les métas-données des objets et les profils, seuls les objets dont les attributs répondent au profil de l'utilisateur seront sélectionnés.

La figure 2.2 suivante illustre le principe du filtrage d'information :



*Figure 2.2. Modèle du Filtrage d'Informations.*

### 11.7. Types des systèmes du filtrage d'information

La fourniture du contenu personnalisé est souvent considérée comme une tâche du filtrage ou de recommandation. Le filtrage peut être vue comme un problème de prédiction, où le système doit essayer de prévoir la pertinence d'un objet à un utilisateur par l'emploi de différentes techniques et stratégies. Généralement les systèmes de recommandation diffèrent par des types de données décrivant les utilisateurs et les objets ; ou par des algorithmes utilisés pour effectuer la prédiction [Mobasher,2006].

D'un point de vue architectural et algorithmique, Mobasher a distingué dans [Mobasher,

2006] trois catégories de systèmes de recommandation : systèmes du filtrage à base de règles ; systèmes du filtrage par contenu ; systèmes du filtrage collaboratif. Notons que l'objectif de ces approches est de recommander les objets susceptibles intéressé par un utilisateur.

### **II.7.1. Systèmes du filtrage à base de règles**

Ce type de systèmes contient des règles de décision générées manuellement ou automatiquement pour recommander des objets aux utilisateurs. Plusieurs sites Web notamment du e-commerce intégrant les technologies de recommandation en utilisant des systèmes à base de règles générées manuellement. Certains systèmes permettent aux administrateurs de spécifier les règles, souvent en basant sur les caractéristiques des utilisateurs. Les règles sont utilisées pour affecter le contenu aux utilisateurs qui ont des profils satisfaisant une ou plusieurs conditions de la règle. Des concepteurs de ces systèmes de recommandation fondent beaucoup sur l'ingénierie des connaissances pour construire une base de règles selon des spécificités du domaine. Quelques recherches sont centrées sur les techniques de l'apprentissage automatique pour classifier des utilisateurs dans différentes catégories en basant sur leurs caractéristiques stockées dans le modèle utilisateur. Par conséquent le système dérive automatiquement les règles de décision qui peuvent être employées pour effectuer la recommandation aux groupes des utilisateurs [Mobasher, 2006].

### **II.7.2. Limites du filtrage à base de règles**

Au-delà de problème habituel du goulot d'étranglement d'ingénierie cognitive, les inconvénients primaires du filtrage à base de règles émanent à partir des méthodes employées pour générer des profils des utilisateurs. L'entrée est habituellement la description subjective des utilisateurs ou de leurs intérêts par les utilisateurs eux-mêmes.

En outre, les profils sont souvent statiques, donc l'efficacité du système commence à dégrader après la fin de durée de validité des profils.

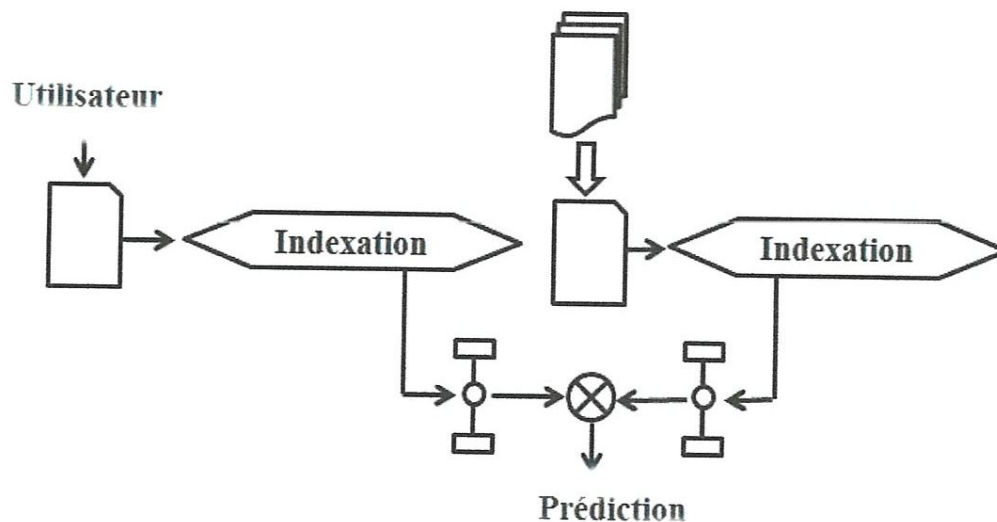
### **II.7.3. Systèmes du filtrage par contenu**

Le filtrage par contenu connu aussi sous le nom de filtrage cognitif, est l'approche la plus anciennement utilisée dans le domaine du filtrage, elle trouve ses racines dans le monde de la recherche d'information. Les objets entrants sont indexés par des termes et comparés avec les profils des utilisateurs représentés sous forme d'un ensemble d'attributs décrivant ses centres d'intérêt principalement. Les fonctionnalités principales du filtrage par contenu visent à sélectionner des objets pertinents par rapport au profil et à mettre à jour ce dernier en fonction



du retour de pertinence émis par l'utilisateur sur les objets qu'il a reçu [Maltz, 1995]. La mise à jour se fera par l'insertion des nouveaux attributs abordés par les objets jugés pertinents. Dans ce type de systèmes, chaque utilisateur a un comportement indépendant des autres comme s'il était entrain de faire une recherche d'information classique [Gallardo, 2005]. En outre, la tâche de génération de recommandation implique habituellement la comparaison des caractéristiques extraites des objets avec des attributs du profil d'utilisateur. Les objets qui sont considérés suffisamment semblables au profil d'utilisateur sont recommandés à l'utilisateur. Particulièrement dans la plupart des systèmes du filtrage par contenu utilisés en Web notamment dans des environnements du e-learning, les caractéristiques du contenu sont les attributs extraits à partir des descriptions des objets pédagogiques. Ces systèmes se fondent souvent sur des normes et spécification de modélisation des objets d'apprentissage tels que : Dublin Core, ARIADNE, LOM, ...etc. et également sur des standards de modélisation des utilisateurs tels que : IMS LIP, PAPI, eduPerson, vCard, EDS, modèle de Chen et Mizoguchi, Dolog LP, FOAF, ...etc.

La (figure 2.3) ci-dessous illustre le principe de filtrage par contenu :



*Figure 2.3. Filtrage par contenu.*

#### II.7.4. Limites du filtrage par contenu

Le filtrage par contenu souffre de certaines limitations, parmi les quelles, citons principalement :

- difficulté d'indexer des objets multimédias (image, son, vidéo), ce qui constitue une grande restriction et un véritable obstacle pour l'application de cette technique ;

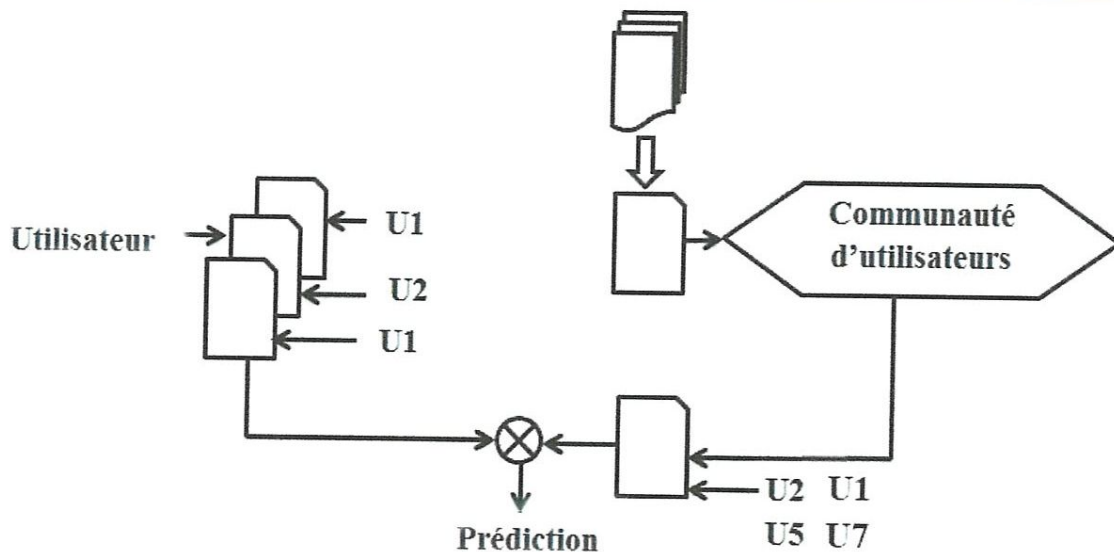


- pas d'intégration d'autres critères de pertinence à la description des objets ou au profil d'utilisateur tels que : la qualité scientifique des objets présentés, la fiabilité de la source d'information, le degré de précision des objets présentés,...etc.
- les systèmes actuels ne permettent pas, dans certains cas, de filtrer tous les objets surtout ceux qui sont proches thématiquement mais décrits différemment. Cependant pour remédier à ce problème, une des solutions consiste à intégrer manuellement les nouveaux thèmes dans le profil de l'utilisateur.
- performance et efficacité faibles des mécanismes de comparaison entre les attributs des objets et des utilisateurs.

### II.7.5. Systèmes du filtrage collaboratif

Le travail collaboratif suscite de jour en jour l'intérêt de compagnie socio-économique et même le monde de l'université et la recherche scientifique, il n'y a qu'à voir le nombre croissant de plateformes et d'outils qui sont développés dans ce domaine. Ainsi, dans la plupart des communautés partageant des centres d'intérêt, les personnes se recommandent ou s'échangent régulièrement des objets avec leurs amis et collègues[Gallardo, 2005]. Le filtrage collaboratif connu aussi sous le nom de filtrage social, vient renforcer l'idée que les personnes à la recherche d'informations devraient pouvoir se servir de ce que d'autres ont déjà trouvé et évalué [Maltz, 1995]. La différence entre le filtrage par contenu et le filtrage collaboratif est que le filtrage par contenu est appliqué pour un seul utilisateur, par contre le filtrage collaboratif est appliqué à une communauté d'utilisateur. Le principe du filtrage collaboratif est basé sur les décisions des autres utilisateurs, le système recommande à un utilisateur des objets préférés par d'autres utilisateurs qui ont des profils proches de lui, ç-à-dire qui appartient à la même communauté ; et des objets populaires ç-à-dire qui sont préférés par la majorité des utilisateurs. Par conséquent, au lieu de calculer la similarité entre le profil et l'objet, cette approche calcule la similarité seulement entre les profils des utilisateurs.

La figure 2.4 ci-dessous illustre le principe du filtrage collaboratif :



*Figure 2.4. Filtrage collaboratif.*

### II.7.6. Limites du filtrage collaboratif

Le filtrage collaboratif souffre de certaines limitations, parmi les quelles, citons principalement :

- Le démarrage à froid qui est un grand problème pour le filtrage collaboratif, cela signifie qu'au démarrage du système, il n'y a pas encore d'évaluations disponibles pour que le système puisse donner des recommandations. Parce que le système doit avoir un minimum d'évaluations des utilisateurs pour filtrer efficacement des objets pour un utilisateur et que le calcul de prédiction prend tout son importance or les profils au départ sont vides et une période d'apprentissage plus au moins longue est toute fois nécessaire pour les constituer, la période durant laquelle beaucoup d'utilisateurs peuvent abandonner d'utiliser le système [Berrut, 2003] ;
- Dans ce type du filtrage, le système dépend fortement de son utilisation et du comportement de l'ensemble des utilisateurs en terme d'évaluations des objets. Le problème des nouveaux objets qu'ils n'ont pas encore été évalués donc ils ne peuvent pas être diffusés aux utilisateurs concernés;
- Les utilisateurs qui ont des centres d'intérêts singuliers, risquent de ne pas recevoir des propositions et des recommandations par le système. Par conséquent, ils se retrouveront certainement isolés des autres utilisateurs du système ;
- Si le nombre des utilisateurs est trop petit par rapport au nombre d'objets dans le système, alors qu'il n'y a pas assez de données pour faire une prédiction. C'est pour cette raison, quelques systèmes de recommandation ont choisi d'utiliser autre type de filtrage.



### **II.7.7. Filtrage hybride**

Les systèmes de recommandation hybrides combinent les techniques de recommandation (deux ou plus) pour gagner une meilleure performance en tirant profit des avantages des approches employées, et limitant les problèmes qui leur sont liés.

La combinaison la plus fréquente se fait entre le filtrage collaboratif et une autre technique en essayant d'éviter le problème de nouvel utilisateur et/ou item. En général, cette autre technique consiste en le filtrage basé sur le contenu tel que: les profils sont orientés contenu, et la comparaison entre ces profils donne lieu à la formation de communautés permettant le filtrage collaboratif. Autrement dit, un item peut être bien apprécié soit par l'utilisateur lui-même, soit par un ensemble d'utilisateurs qui représentent des profils similaires.

Des résultats empiriques démontrent que les recommandations hybrides sont plus efficaces que dans le cadre des approches dites 'pures'. L'approche hybride est l'approche la plus communément adoptée ces dernières années.

## **II.8. Calcul de prédiction**

La prédiction est l'opération qui consiste à estimer l'intérêt que portera un utilisateur pour un objet qui serait recommandé pour lui [Maltz, 1995].

### **II.8.1. Prédiction dans le filtrage cognitif**

Dans le cas du filtrage par contenu, la prédiction de l'opinion de l'utilisateur sur un objet donné, est estimée par un calcul de proximité entre les attributs du profil de l'utilisateur et les attributs de description de l'objet. Ce calcul de similarité peut se faire en utilisant plusieurs techniques, tels les techniques de l'intelligence artificielle (réseaux de neurones, arbre de décision, support vector machine,...etc.) ; ou l'approche booléenne ; ou l'approche vectorielle ; ou l'approche de LSI ;...etc. [Nouali, 2006].

### **II.8.2. Prédiction dans le filtrage collaboratif**

En filtrage collaboratif, la prédiction de l'opinion de l'utilisateur sur un objet se calcule par rapprochement entre les évaluations passées de l'utilisateur et des autres utilisateurs d'une communauté sur les mêmes objets. Le filtrage collaboratif tient compte de la proximité entre les utilisateurs, afin de ne recommander un objet qu'aux utilisateurs proches de celui qui l'a



apprécié [Lee, 2005]. Une spécificité du filtrage collaboratif réside dans le fait que cette proximité qui contribue à la formation des communautés d'utilisateurs est estimée sur l'unique base de la liste de leurs évaluations passées sans tenir aucun compte du contenu des objets. Les systèmes de filtrage collaboratif peuvent être classés en trois grandes catégories : les algorithmes à base de « mémoire », les algorithmes « d'apprentissage en ligne » et les algorithmes à base de « modèle ».

#### **II.8.2.1. Algorithmes à base de « mémoire »**

Ce type d'algorithmes consiste à faire la prédiction à partir d'informations partielles concernant l'utilisateur actif et un ensemble de poids calculés à partir de la base de données des évaluations des utilisateurs. Selon la façon du calcul des poids, ces algorithmes se discriminent en « algorithme basé sur la corrélation » et « algorithme basé sur la similarité de vecteurs (Cosinus) ».

Ces algorithmes sont largement employés dans le domaine du e-commerce dont l'émergence a permis le développement de nombreux systèmes de recommandation pour filtrer un ensemble d'articles pouvant intéresser les clients, pour cela on applique ce type d'algorithmes qui sont les plus utilisés jusqu'à maintenant pour le calcul de prédiction dans les systèmes de recommandations commerciales.

Cependant ces algorithmes deviennent compliqués à déployer avec l'augmentation linéaire du nombre de clients [Ouraiba, 2008].

#### **II.8.2.2. Algorithmes d'apprentissage en ligne**

Ce type d'algorithmes consiste à utiliser un agent de prédiction pour chaque utilisateur.

Chaque agent sera confronté à un ensemble d'essais avec une prédiction à faire à chaque étape en fonction des agents du même environnement qui pourront avoir un comportement similaire, neutre ou opposé à la fonction laquelle cet agent cherche à atteindre. Cependant ces algorithmes deviennent compliqués à déployer avec l'augmentation du nombre des utilisateurs et par conséquent le nombre des agents.

Ceci nécessite la création d'un modèle prédictif indépendant de nombre des utilisateurs [Berrut, 2003].

#### **II.8.2.3. Algorithmes à base de « modèle »**

Ce type d'algorithmes consiste à construire un modèle prédictif à partir de la base de données des évaluations des utilisateurs. Une prédiction est effectuée par le calcul de probabilité d'une

valeur souhaitée d'une évaluation. Ce calcul est opéré par différents modèles, tels que le modèle à base de clusters (modèle probabiliste) : EM, Modèle Hiérarchique...etc. ; le modèle à base de réseau bayésien ; et toutes autres techniques de l'intelligence artificielle [Berrut, 2003].

## **II.9. Une note sur l'évaluation des systèmes de recommandation**

Les systèmes de recommandation posent quelques problèmes particuliers en leur évaluation d'efficacité. Il est très difficile de définir exactement leur comportement désiré, car le « meilleur » comportement est différent pour différents utilisateurs, voire pour le même utilisateur aux différents moments ou contextes. D'ailleurs, le comportement du système dépendra de la qualité des informations stockées dans le profil d'utilisateur. Si le système a des informations insuffisantes ou incorrectes sur l'utilisateur, alors il ne peut pas interagir d'une manière "idéale" avec l'utilisateur. Ces facteurs rendent généralement très difficile de faire des expériences artificielles de mesure de performance et d'efficacité d'un système de recommandation. Pour faire une évaluation significative, le système à évaluer doit être employé par des vrais utilisateurs avec des vraies caractéristiques (but, préférences,...etc.). Le succès peut alors être jugé de certaines manières : la satisfaction d'utilisateur par le système et par ses recommandations est le critère primaire du succès. Ceci peut être mesuré en utilisant des questionnaires remplis par des utilisateurs.

De plus, pour les systèmes du e-learning en particulier, on trouve un autre critère du succès qui consiste à mesurer "le degré de compréhension de la matière enseignée" par l'évaluation des connaissances de l'apprenant après l'utilisation du système (ceux-ci pourraient être formellement ou informellement). Dans ce cas, un certain système de "contrôle" devrait également être utilisé, pour permettre la mesure de "la valeur ajoutée" par le filtrage.

## **II.10. Conclusion**

Dans ce deuxième chapitre, nous avons présenté en détail les systèmes de recommandation qui ont montré leur preuve en termes d'offre efficace d'objets pertinents aux utilisateurs en générale et aux apprenants en particulier. Nous nous sommes intéressés au concept du filtrage. Nous avons défini ce que c'est le filtrage. Il est d'un grand secours pour un utilisateur. Il lui offre le contenu qui lui est sur mesure.

Alors, après la présentation d'une brève historique et des domaines d'application des systèmes de recommandation, nous avons montré ses caractéristiques et le mécanisme de leur fonctionnement. Ensuite, nous avons étalé les trois types de ces systèmes du filtrage : à base de règles, par contenu, collaboratif, et hybride. En issus, nous avons montré la façon du calcul de prédiction. En dernière, nous finissons par une note sur l'évaluation des systèmes de recommandation.



# Chapitre 3

## Etude conceptuelle

---

## Chapitre 3 : Etude conceptuelle (la technique CSHTRL)

### III.1. Introduction

Les systèmes de recommandation sont employés récemment dans les systèmes interactifs comprenant l'apprentissage en ligne. Malgré qu'ils ont prouvés largement leur puissance dans plusieurs domaines, le e-Learning a introduit cette technologie récemment. Néanmoins, ils ont montré leurs avantages dans ce secteur en particulier aux apprenants.

② / À travers ce chapitre qui constitue le noyau de ce mémoire, nous allons montrer les contributions des systèmes de recommandation dans l'e-Learning en utilisant la technique CSHTRL, est une technique de recommandation personnalisée. Nous débuterons par les motivations qui nous ont orientées vers la réalisation de ce travail, après nous décriront étape par étape cette technique.

### III.2. Motivations et Objectif

- Les environnements des E-Learning estiment que tous les apprenants sont similaires, de sorte que ces environnements ne sont pas en mesure de répondre aux besoins des apprenants, car ces apprenants sont différents en fonction de leurs niveaux de connaissance, leurs progrès d'apprentissage, leurs capacités intellectuelles, et ainsi de suite, donc tout le monde préfère des contenus qui correspondent à ses caractéristiques.
- L'environnement d'apprentissage doit donner plus d'importance à l'apprenant et à devenir de plus en plus personnalisé en introduisant les caractéristiques cognitives qui sont liées à l'apprenant.

La personnalisation dans e-Learning consiste à proposer des contenus pédagogiques appropriés aux apprenants, prenant en compte leur caractère moral. Le système de recommandation est l'outil le plus utilisé pour la personnalisation.

Notre objectif vise à développer un environnement d'apprentissage personnalisé pour pouvoir recommander des cours les plus convenables aux besoins et préférences d'apprenant afin de lui faciliter l'apprentissage et nous nous sommes concentrés sur la personnalisation des contenus éducatifs en utilisant les systèmes de recommandation.

Pour cela on a utilisé la technique CSHTRL (Cold Start Hybrid Recommend Taxonomy For e-Learning).

Cette technique a été proposée pour offrir aux apprenants des contenus pédagogiques personnalisés, et pallier aux problèmes de démarrage à froid en se basant sur les taxonomies

### III.3. CSHTR (l'origine de CSHTRL)

CSHTR (Cold Start Hybrid Taxonomy Recommender) est l'une des approches existantes qui utilisent la taxonomie. L'idée dans le système CSHTR est de comparer les préférences taxonomiques d'un utilisateur qui se trouve dans une situation de démarrage à froid aux préférences taxonomiques des différents groupes et de l'affecter à celui qui est jugé le plus proche puis la recommandation sera basée sur les préférences de ce groupe. Les regroupements dans ce système sont réalisés d'une manière naturelle à la base des évaluations qu'ont faites les utilisateurs.

On se réfère, pour la description de cette technique, à son concepteur Li-Tung Weng.

CSHTR principalement pour résoudre le problème de démarrage à froid relatif à un utilisateur et/ou un nouvel item en se basant sur une approche hybride qui utilise le concept des taxonomies.

#### III.3.1. Objectifs de CSHTR

CSHTR utilise le FC basé item en plus de l'information taxonomique sur les items pour générer la recommandation. Cette information est une très simplifiée d'une connaissance, mais aussi, elle peut être considérée comme une représentation du contenu des items du point de vue que les profils des items contenant cette information au lieu d'un ensemble de mots clés.

#### III.3.2. Description de la technique

##### III.3.2.1. La représentation du monde visé

On envisage un monde avec :

- Un ensemble d'utilisateurs  $U = \{u_1, u_2, \dots, u_n\}$  ;
- Un ensemble d'items  $I = \{i_1, i_2, \dots, i_m\}$  ;
- Chaque utilisateur  $u \in U$  est associé à un ensemble d'évaluations qu'il a fait sur les items. En basant sur les différentes méthodes d'évaluation, cet ensemble peut être divisé en deux parties :



Formellement, la préférence générale de l'item  $t_i \in T$  dans un groupe  $uc \in UC$  peut être calculé par :

$$cpref(uc, t_i) = \beta \times \psi(uc, t_i) + (\beta - 1) \times \sigma(uc, t_i) \quad (4)$$

Pour prédire la préférence d'un item par un utilisateur, CSHTR vérifie d'abord si cet item est préféré dans son groupe, puis il vérifie la similarité taxonomique entre l'utilisateur et l'item.

### III.3.3 Critiques

#### III.3.3.1. Avantages

Comme le résultat de l'expérimentation le dit, la technique CSHTR a fait épreuve d'efficacité dans les situations de démarrage à froid. Une autre fois le filtrage collaboratif démontre sa puissance, parce que l'information taxonomique est exploitée dans les deux techniques mises en comparaison (i.e. TPR et CSIITR). Les approches hybrides donnent souvent des meilleurs résultats par rapport aux autres, c'est pour ça l'hybridation elle-même constitue souvent un avantage.

#### III.3.3.2 Inconvénients

Malgré que cette technique traite bien les aspects du problème de démarrage à froid relatifs à un nouvel utilisateur et/ou item, mais elle ignore où le problème est global ; c'est-à-dire le démarrage à froid couvre l'état de tous les objets dans le système (i.e. utilisateur/items), ce qui entraîne que même les groupes ne peuvent pas être construit à cause du manque des évaluations communes entre les utilisateurs.

Dans ce cas, la technique CSHTR ne pourra pas fonctionner, car elle suppose l'existence basique des groupes dans le système, où un ensemble de données (i.e. d'évaluation) suffisent pour faire le regroupement des utilisateurs.

### III.4. Présentation de CSHTRL [Benhamdi, 2011]

CSHTRL (Cold Start Hybrid Taxonomy Recommender for e learning) est une technique de recommandation personnalisée elle a été adaptée pour le domaine de l'e-Learning, c'est une version améliorée de la technique de recommandation CSHTR elle se réfère à une situation plus sévère, où de nouveaux apprenants n'ont pas le temps pour lire et évaluer les documents.

À cause de cette raison la technique (CSHTRL) propose un questionnaire aux apprenants, ce questionnaire se compose de trois parties, la première partie pour définir le domaine d'intérêt de l'apprenant, la deuxième partie pour définir le niveau éducatif de l'apprenant, et la troisième partie pour mesurer la capacité mémoire de l'apprenant. Pour mesurer la capacité mémoire d'un apprenant on a utilisé une méthode de psychocognitive appelé ISR[] (Immediate Serial Recall) par B. M. Kim and Q. Li, adopté pour l'évaluation des patients atteints des lésions cérébrales, cette méthode a été développée par Daneman et Carpenter, comprend généralement la présentation des séries de phrases, les apprenants lisent ces phrases et après une seule présentation ils doivent les répéter dans l'ordre correct, les articles peuvent inclure des mots, des chiffres ou des lettres.

Donc, l'idée générale consiste à remplacer les évaluations des apprenants adoptées par CSHTR, par les poids des termes choisis par l'apprenant, le background des connaissances de l'apprenant, et la capacité de stockage de l'information (capacité de mémoire), à partir du questionnaire proposé, afin de faciliter la tâche et lui décharge de lire et évaluer les items.

### III.4.1. Description formelle

#### III.4.1.1 Profil apprenant

La Recommandation ne se rapporte pas uniquement au contenu des ressources, mais également aux attributs descriptifs des apprenants. Donc, le profil de l'apprenant doit être constitué.

#### ➤ construction des vecteurs taxonomiques :

Dans la technique CSHTR, l'utilisation de l'information taxonomique permet de représenter sous forme d'un vecteur les préférences des utilisateurs. Une conversion de leurs évaluations à des préférences taxonomiques est une technique qui a été implémentée par Zeigler.

Dans la technique CSHTRL, après avoir répondu à les questions du questionnaire (testé) proposé on obtient le vecteur taxonomique de l'apprenant, et le vecteur taxonomique est noté comme suit :

$$\vec{v} = (s_1, s_2, s_3, s_{t1}, s_{t2}, s_{t3}, s_{t4}, s_{t5}, t_{s1}, t_{s2}, t_{s3}, t_{s4}, t_{s5}, mc)$$

Si : le score du test numéro i.

Ti : le temps du test numéro i.

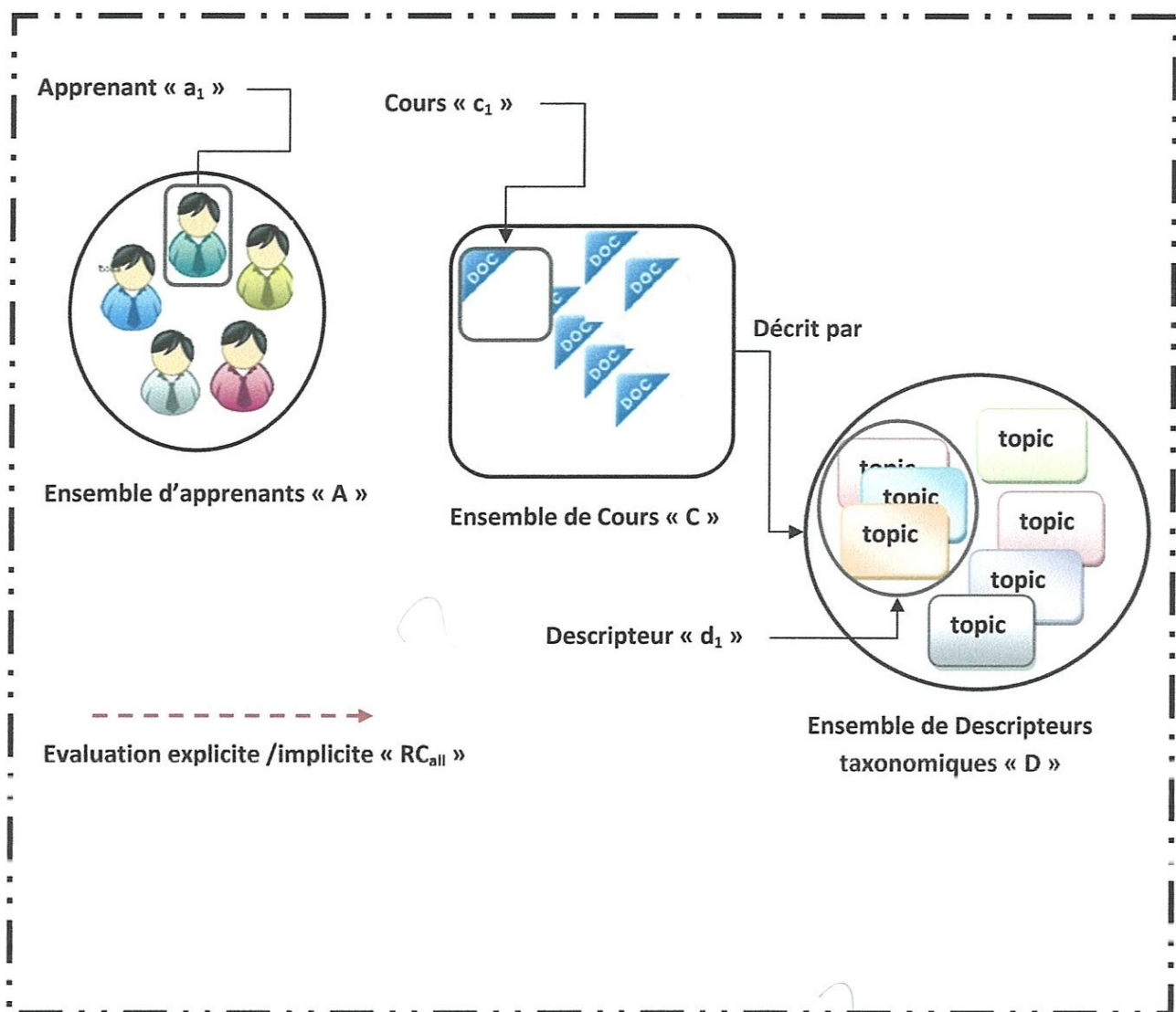
mc : la capacité mémoire de l'apprenant.

### III.4.1.2 Description de l'environnement

Nous décrivons la présentation de la technique CSHTRL. Et nous utilisons des simples notations pour faciliter sa compréhension.

- un ensemble d'apprenants  $A = \{a_1, a_2, \dots, a_C\}$ .
- un ensemble de cours  $T = \{t_1, t_2, \dots, t_k\}$ .
- chaque élément  $T$  est représenté par un ensemble de descripteurs.
- $D(t) = \{d_1, d_2, \dots, d_L\}$ . Un descripteur est une séquence de rubriques de la taxonomie commandées, noté  $d = \{p_0, p_1, \dots, p_q\}, j \in D(t), t \in T$ .

Pour mieux comprendre, récapitulons ce que nous venons de dire comme suit :

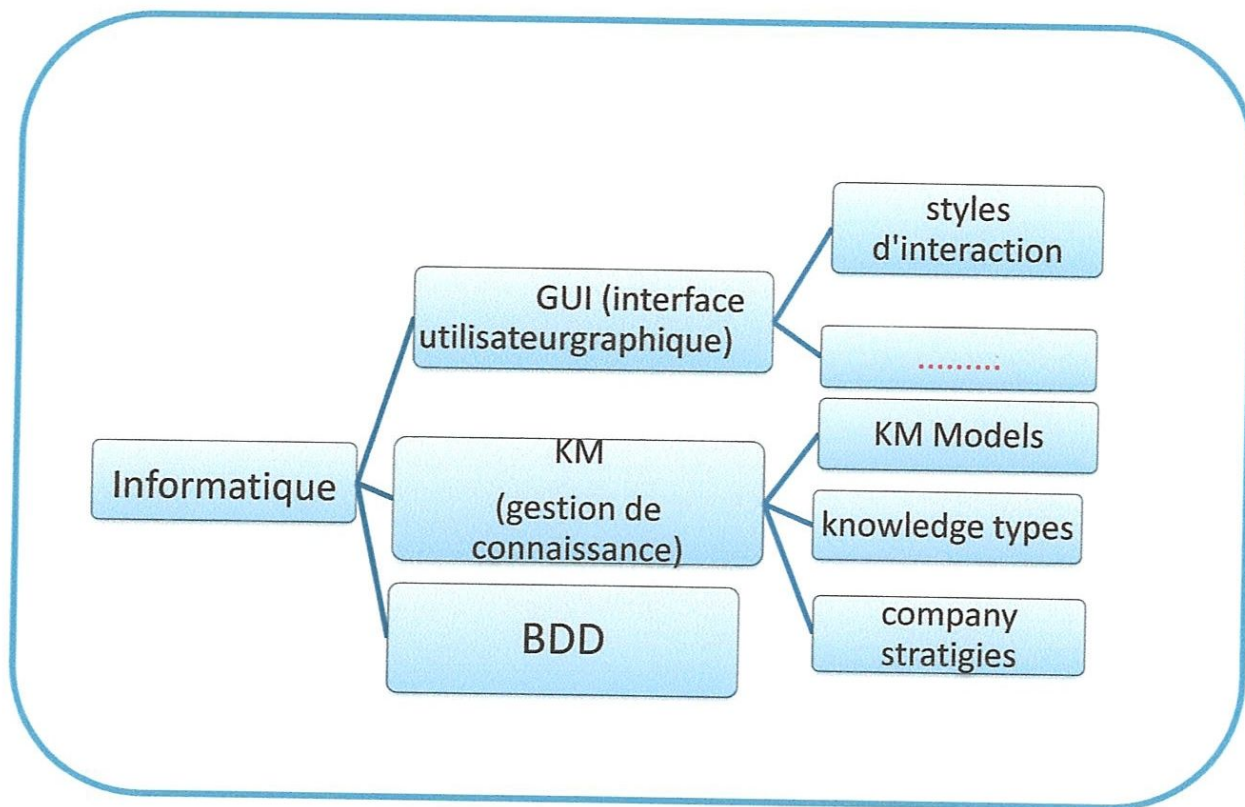


*Figure 3.1. Description de l'environnement*

Il faut noter que cet ensemble de descripteurs se diffère de celui de e-commerce.



Pour mieux comprendre, un exemple de l'information taxonomique que nous avons définie est comme suit :



*Figure 3.2. Descripteur taxonomique en e-Learning*

➤ **Le regroupement :**

Afin de construire des groupes d'utilisateurs, Weng employé la mesure de corrélation pour calculer la similarité des préférences pointe entre deux utilisateurs :

$$sim(a, i) = \frac{\sum (r_{a,j} - \bar{r}_a)(r_{i,j} - \bar{r}_i)}{\sqrt{\sum_{j \in I} (r_{a,j} - \bar{r}_a)^2 (r_{i,j} - \bar{r}_i)^2}} \quad (5)$$

L'idée pour calculer la similarité dans CSHTRL ce n'est pas. Seulement le calcul de la similarité entre les apprenants, mais aussi elles se basent sur leurs vecteurs taxonomiques, et le temps qu'ils passent à différents niveaux du test, sachant que nous attribuons à chaque test un poids plus important que le précédent, la nouvelle mesure de similarité entre les utilisateurs x et y, peut être établie comme défini en (6). Le premier terme de l'équation se réfère à la

différence des scores des niveaux du test ( $Cx$ ) de l'étudiant  $x$ , et les scores niveaux du test ( $Cy$ ) de l'étudiant  $y$ , tandis que le second terme se réfère à la similarité des apprenants en terme des temps de réponse, le troisième terme se réfère à la similarité traditionnelle (similarité de Pearson qui a été utilisée dans la technique CSCTRL), entre les apprenants en fonction de leurs préférences taxonomiques et en appliquant la corrélation métrique, et le terme de suite représente la similarité des capacités mémoire entre les apprenants  $mc_x - mc_y$ .

Ou  $mc$  est calculée par l'expression (8)

$$new\_sim(x, y) = \sum_{n=1}^L n\alpha(Cx_n - Cy_n) + \frac{1}{\varphi(tx, ty)} + sim(x, y) + mc_x - mc_y \quad (6)$$

$$\alpha \in [0..1]$$

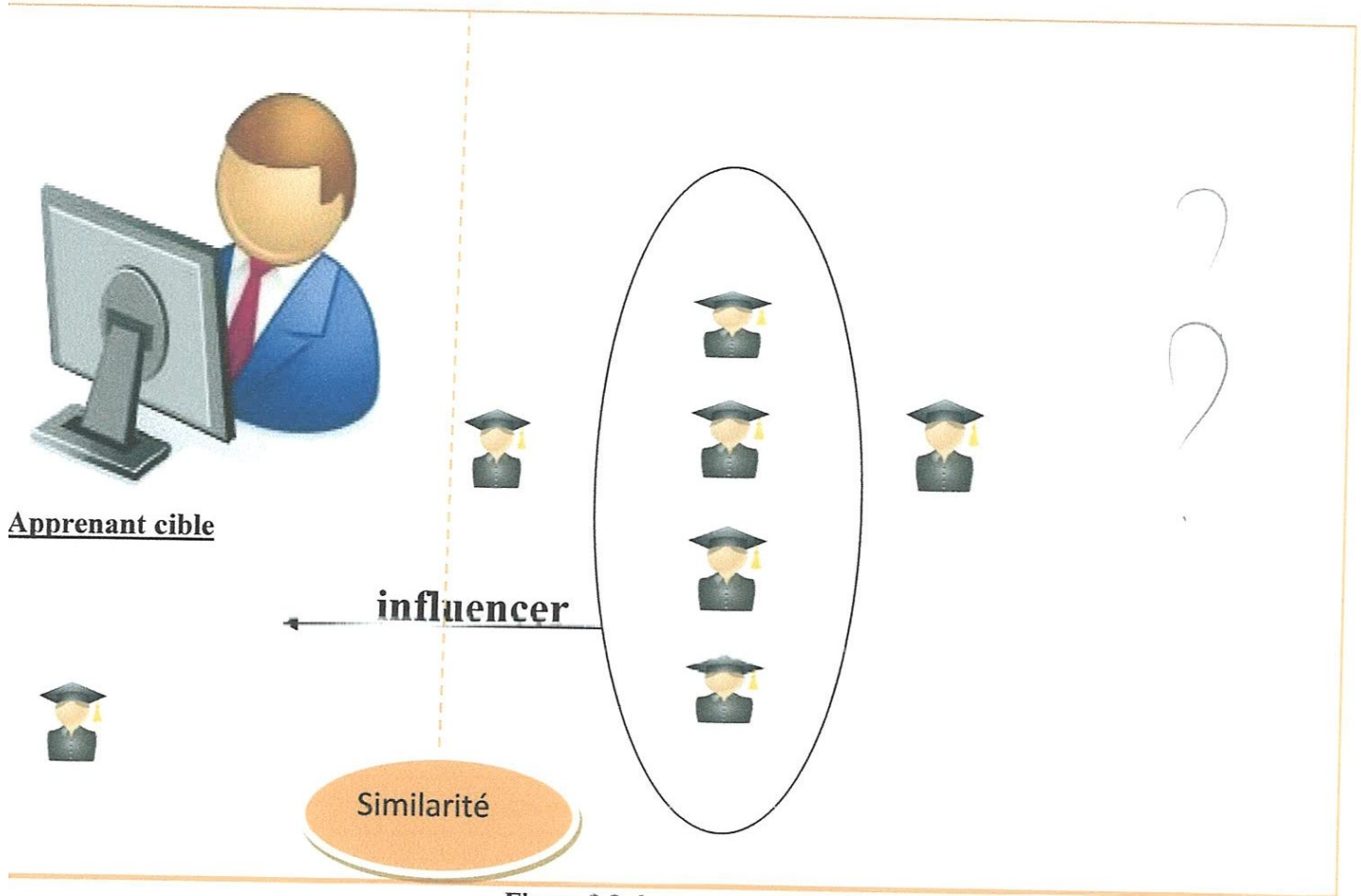
$$\varphi(tx, ty) = \frac{\sum_{z=1}^L |tx_z - ty_z|}{L} \quad (7)$$

$$mc_x = \frac{l}{g} \quad (8)$$

$g$ : est le nombre de tous les termes qui composent la longue liste présentée aux apprenants.

$l$ : représente-lui numeros de tous les termes qui composent la longue liste des propositions retenues par les apprenants

$L$ : représente le nombre des tests.



*Figure 3.3. Le regroupement*

➤ **Génération de la recommandation :**

Après l'attribution de l'apprenant au groupe adéquat, il sera normalement recevoir des recommandations fondées sur les articles préférés de son groupe. Selon Weng, les articles préférés peuvent être identifiés par le nombre d'utilisateurs qui ont évalué l'article (popularité) et la moyénno des notes. Mais un article qui a été apprécié par un grand nombre d'apprenants, cela ne signifie pas que cet article est populaire parce que les apprenants peuvent lui assigner des valeurs faibles, de sorte que la technique CSHTRL a proposé pour calculer les articles popularité sur la base des évaluations qui dépassent (ou égale) la moitié de la note maximale (sur l'échelle de 0-1). Formellement, la préférence générale de l'item  $t_i$  dans un groupe  $uc$  peut être calculée par la formule (9) suivante :



$$cpref(uc, t_i) = \beta \times \psi(uc, t_i) + (\beta - 1) \times \sigma(uc, t_i) \quad (9)$$

$0 \leq \beta \leq 1$  est une variable commandée de l'étudiant pour rajuster les poids entre la préférence moyenne de l'apprenant et la popularité de l'item.

$\beta$  Est fixé à 0,7.

Ou :

$$\sigma(uc, t_j) = \frac{\sum_{u \in uc} r(a, t_j)}{|uc|} \quad r(a, t_j) \in [0.5.. 1] \quad (10)$$

$$\psi(uc, t_j) = \frac{\sum_{a \in uc} r(a, t_j)}{|uc(t_j)|} \quad (11)$$

$Uc$  : représente le groupe des apprenants similaires.

$|uc|$  : représente le nombre d'apprenants dans le groupe  $uc$ .

$r(a, t_j)$  : représente l'évaluation de l'apprenant  $a$  dans le groupe  $uc$  sur l'item  $t_j$

$(uc, t_j)$  : se réfèrent à la moyenne de toutes les évaluations sur l'item  $t_j$ .

$|uc(t_j)|$  : désigne le nombre d'étudiants similaires qui ont évalué  $t_j$ .

CSHTRL contrôle si l'article est couramment préféré par les apprenants semblables et calcule la similarité taxonomique entre l'utilisateur cible et l'item afin de prédire la notation de l'utilisateur cible sur un élément donné (pour plus de détails se référer à (12)).

$\varphi$  Est réglé à 0,5.

$$rank_{a,t} = \varphi \times cpref(uc, t) + (1 - \varphi) \times t\_sim(\vec{v}_a, t) \quad (12)$$

L'algorithme suivant résume recommandation étapes définies par boosted\_CSHTRL.

❖ Algorithm cold start taxonomy recommender (a1, k)

Where  $a1 \in A$  is a given target student

$K$  is the number of items to be recommended

- 1) SET  $T$ , the candidate items list
- 2) FOR EACH  $t \in T$
- 3) SET  $uc = t\_cluster(a1)$
- 4) SET  $rank_{a1,t} = rank_{a,t} = \varphi \times cpref(uc, t) + (1 - \varphi) \times t\_sin(\vec{v}_a, \vec{t})$
- 5) END FOR
- 6) Return the top  $k$  items with highest  $rank_{a1,t}$  score to  $a1$ .

Pour mieux comprendre notre système, la figure ci-dessous l'explique :

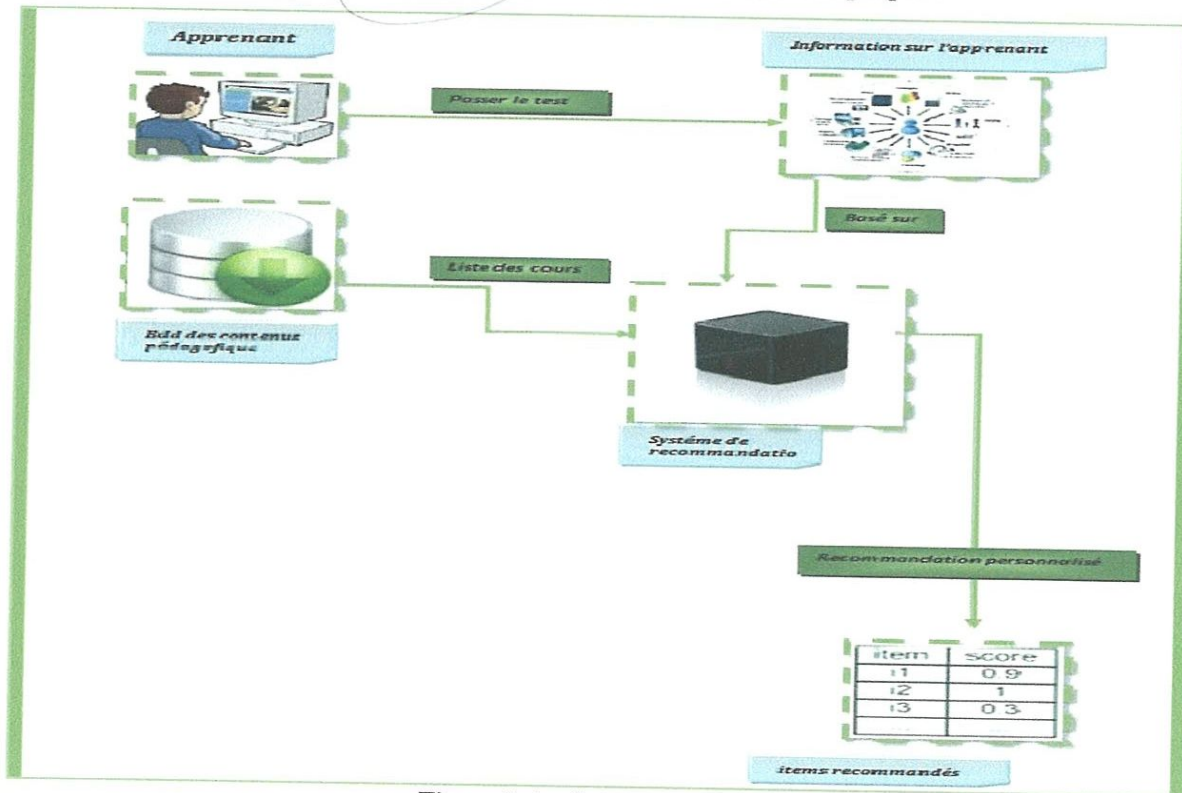


Figure 3.4 : description du système

### III.3 Conclusion

Dans ce chapitre nous avons montré que la nouvelle approche de recommandation CSHTRL est la meilleure extension de CSHTR, ça combine entre les préférences d'items des utilisateurs et les préférences taxonomiques supplémentaires pour faire de meilleures recommandations de qualité ainsi que d'atténuer le problème de démarrage à froid. CSHTRL est spécialement conçu pour être appliqué en e-Learning en donnant beaucoup d'importance aux propriétés des élèves : les connaissances de base, le niveau et leur capacité mémoire, pour faire des meilleures recommandations personnalisées, mais on ne peut juger la puissance de ces choix stratégiques qu'après la phase d'expérimentation dont on va parler dans le chapitre suivant.



# Chapitre 4

## Implémentation

---

## Chapitre 4 : Implémentation

### IV.1. Introduction

L'expérimentation est le moyen qui s'offre à nous pour tester la validité et la pertinence de nos choix.

Après avoir effectué la conception de notre technique CSHTRL, nous arrivons dans cette partie à la validation de cette dernière. , tel que nous allons la valider sur un corpus de données afin de mesurer ses performances. Donc, c'est dans ce chapitre ou nous allons récolter les fruits de nos études.

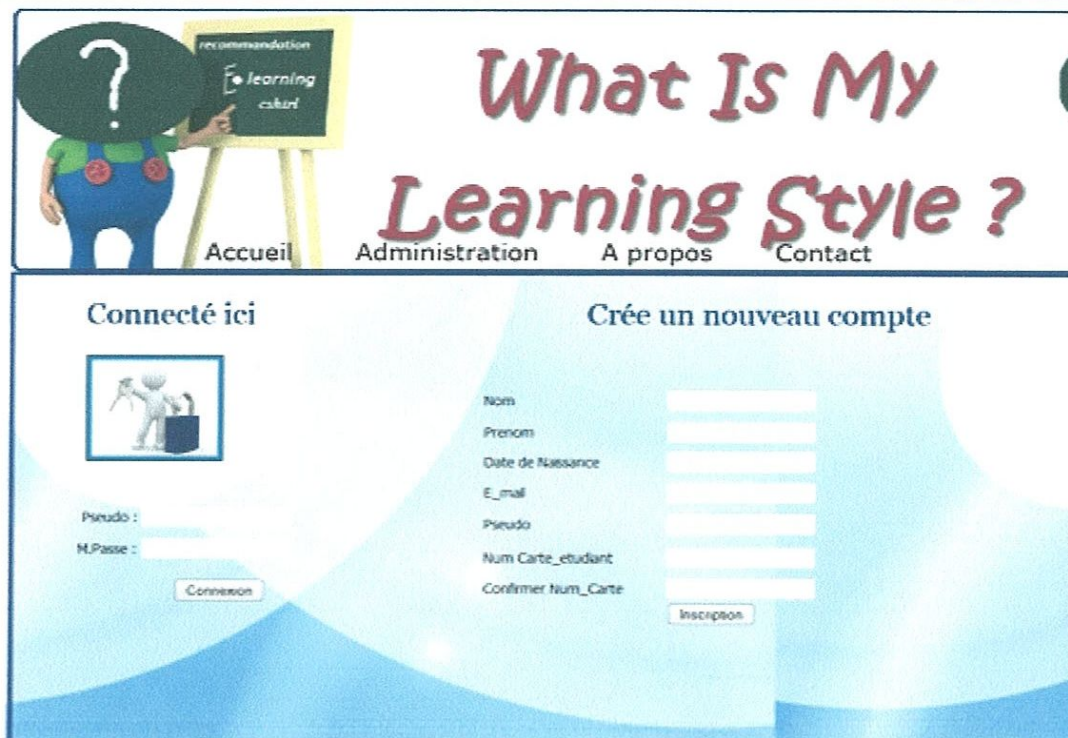
### IV.2. Outils et Implémentation

En effet, le coté implémentation et choix des outils n'est pas un critère de base pour notre travail. L'essentiel pour nous est d'obtenir des bons résultats. Ces résultats ne dépendent réellement que de l'algorithme de recommandation, c'est-à-dire la conception abstraite de notre technique.

Autrement dit, les recommandations qui vont être générées pour un utilisateur donné vont être les mêmes quelques soit le langage de programmation ou l'environnement d'implémentation utilisé.

L'implémentation de notre algorithme ou de n'importe quel algorithme nécessite un langage qui traite bien et facilite la mise en œuvre les formules mathématique, et pour cela nous avons choisis le langage PHP.

Ainsi, nous avons réalisé une démonstration des résultats de la recommandation faite pour des utilisateurs sur un site web pour permettre une meilleure visualisation des résultats, du point de vue que la fonctionnalité de recommandation est populaire et reconnue beaucoup plus dans les sites web tels que Amazon. Un aperçu de cette application et de sa contenu est illustré dans la figure ci-dessous:



***Figure 4.1. Page d'accueil***

En offrant des liens directs vers les autres pages afin de faciliter la navigation. Les liens principaux de cette application sont dans le menu en haut: Accueil, Administrateur, About, Contact.

Cette page d'accueil présente les principes de fonctionnement de notre application, elle comporte deux formulaires : « crée un nouveau compte » et « connecté ici ».

Le formulaire du « crée un nouveau compte » contient 7 champs : nom, prénom, date de naissance, e\_mail, pseudo, Num\_carteetudiant et le dernier confirmer num\_carte. Donc chaque apprenant il faut qu'il possède une carte d'étudiant, après la vérification, le système va générer un mot de passe pour ce nouvel apprenant.

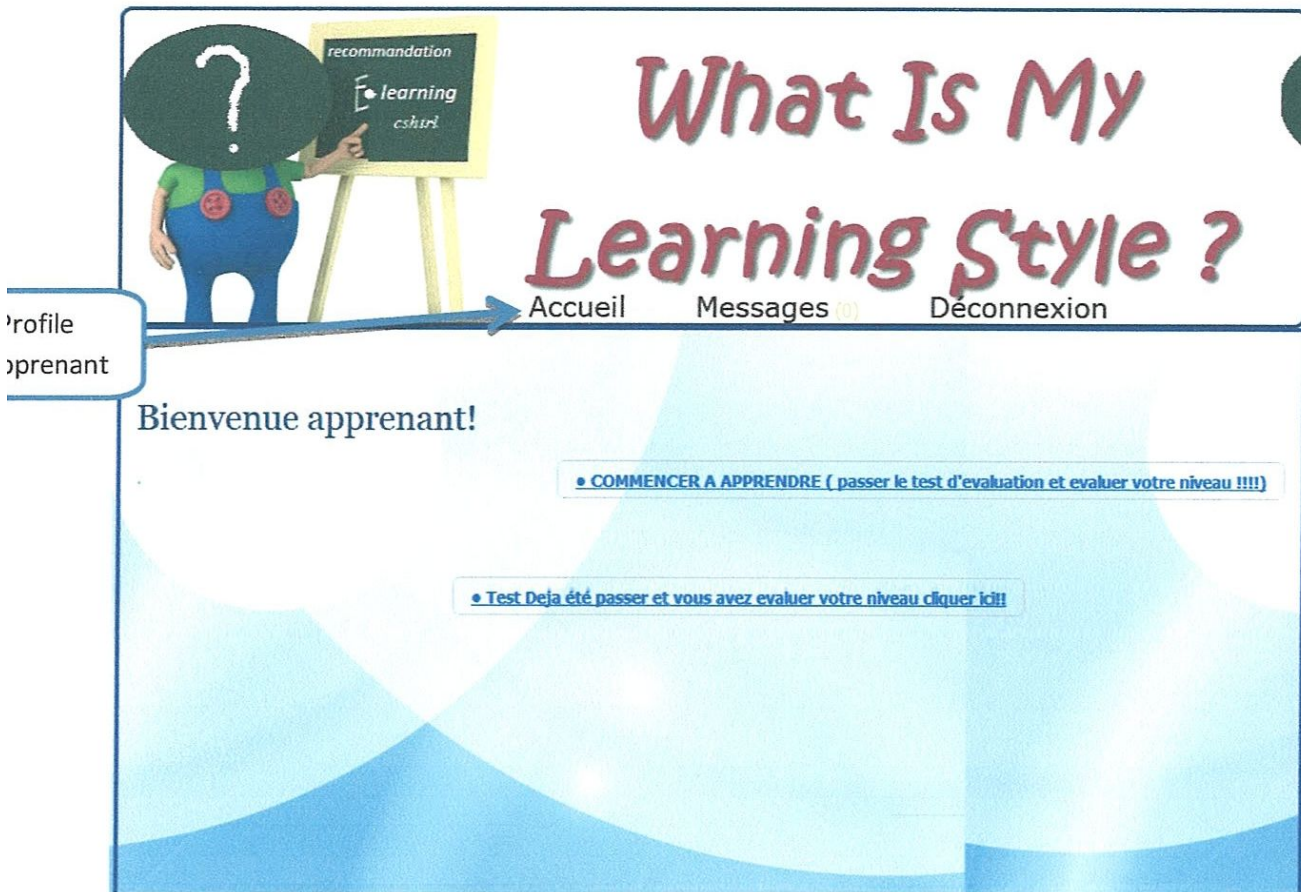
Le formulaire de « connecté ici » contient deux champs textes afin de permettre aux utilisateurs (apprenant/enseignant) de saisir leurs pseudo ainsi que leur mot de passe pour pouvoir accéder à leur compte.

Une fois le mot de passe et le pseudo sont saisis, le système va les vérifier, si ce pseudo n'est pas trouvé dans la base de données, et le mot de passe est faux, le système va afficher un message d'erreur à cet utilisateur.

*Base de Données*



Comme nous montre la figure suivante, une fois que le pseudo et le mot de passe sont reconnus, l'apprenant peut accéder la page suivante :



*Figure 4.2. Compte apprenant 1*

Si l'apprenant entre pour la première fois à son compte doivent cliquer le lien « commencer à apprendre » pour passer le test, si non clique le deuxième lien pour aller à la page d'accueil d'apprenant (profile d'apprenant).

La première étape qui fait dans le test est que l'apprenant doit choisir son domaine d'intérêt (GUI, KM, ou LT). Afin de choisir un Domain et cliquer sur le bouton valider l'apprenant commence le test qui contient cinq niveaux, chaque niveau contient deux questions QCM. Une fois il répond aux questions d'un niveau il clique sur suivant pour passer à l'autre niveau jusqu'au 5ème niveau, une autre page de test de la capacité mémoire, apparais dans (la figure 4.3):



Choisissez les bonnes  
et  
les appropriées réponses

0 minute(s) 25 secondes 9 dixièmes

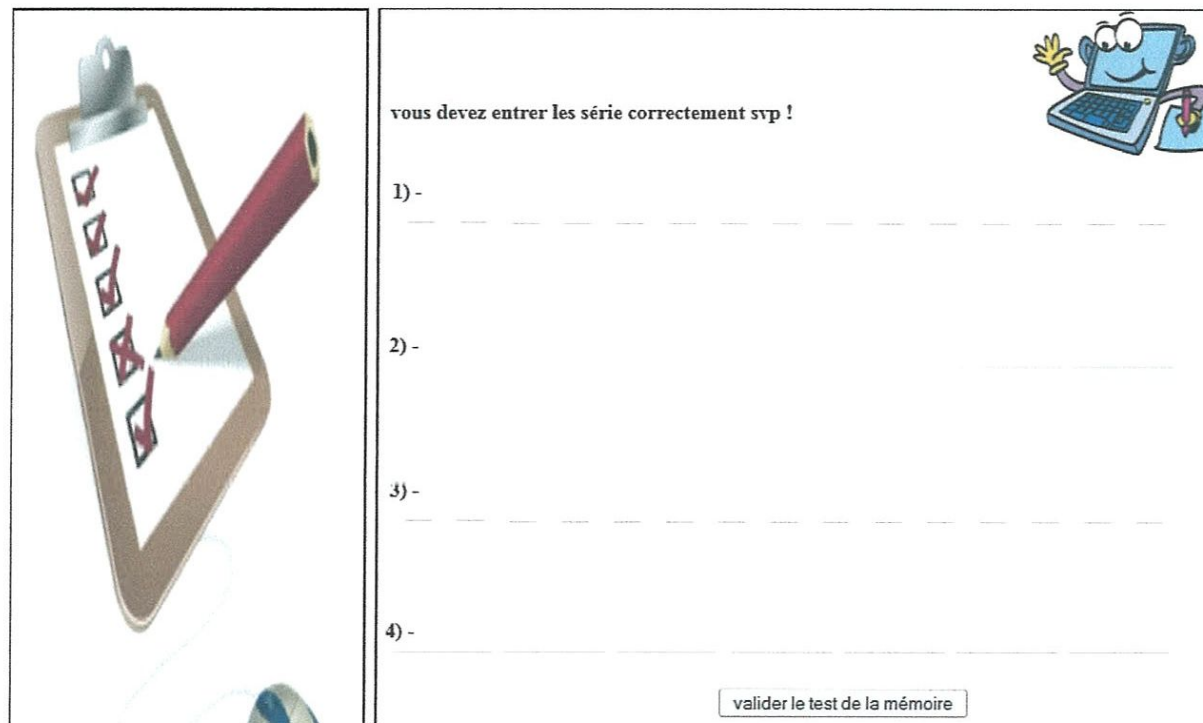
vous devriez vous rappeler ces séries:

- 1)dfghjkkmlbvckkk
- 2)dos-clavier-pascal-java-php-algorithme-table
- 3)235574211059150
- 4)raisin-orange-banan-abricot-ananas-jus-pomme

ok

*Figure 4.3. Test mémoire*

L'apprenant doit mémoriser les séries (mots, lettres, chiffres,...) après une période du temps il clique sur « OK », la page suivante va apparaître :



vous devez entrer les série correctement svp !

- 1)-
- 2)-
- 3)-
- 4)-

valider le test de la mémoire

*Figure 4.4. Test mémoire-2*



Après la validation des tests le système va créer et afficher les recommandations d'apprenant.

### I.V.2.1. Page apprenant

La page d'apprenant ensemble des liens direct vers : passer le test, historique, compte, forum, message, posttest.



*Figure 4.5. Page apprenant*

### I.V.2.2. Page administration

L'administrateur possède plusieurs droits de modification, il peut par exemple supprimer des apprenant, supprimer cour, ajouté ou supprimé des enseignants.

La page d'administration est apparue comme ci-dessous :



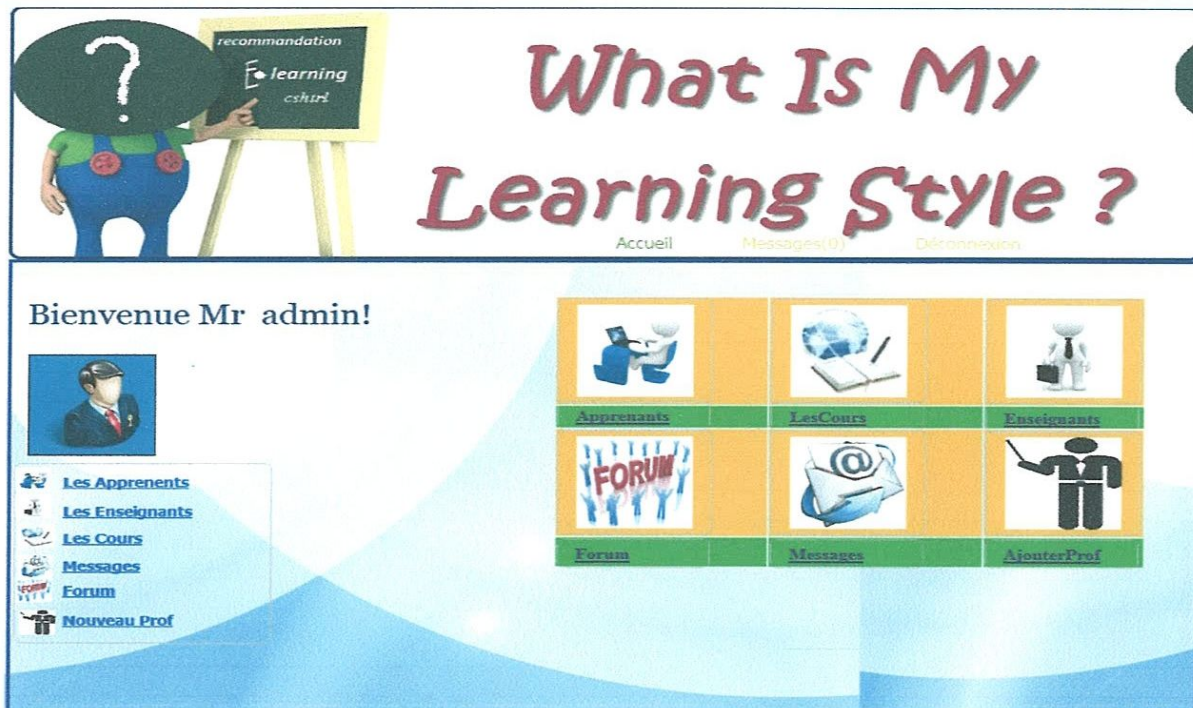


Figure 4.6. Page administration

### I.V.2.3. Page enseignant

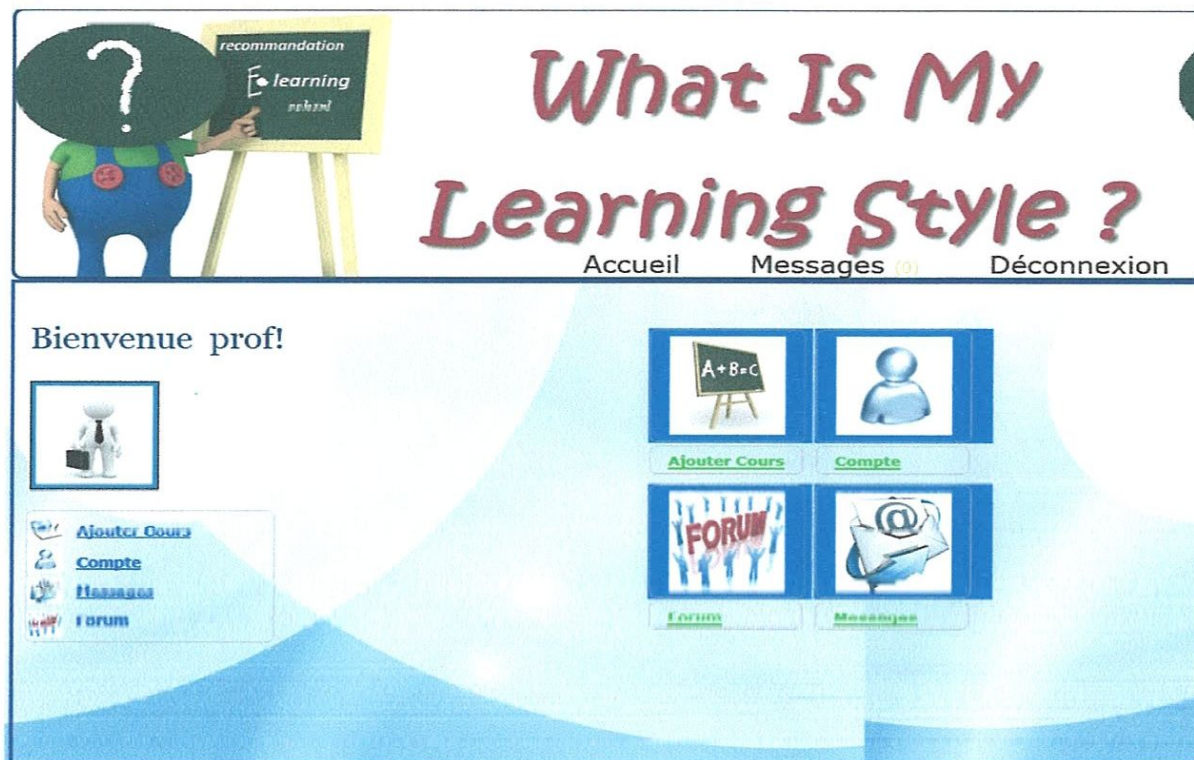


Figure 4.7. Page enseignant

### IV.3. Expérimentation

Nous avons utilisé pour nos expérimentations la même base qui a été utilisé pour valider la technique *Boosted\_CSHTRL*, l'échantillon choisi est composé de 10 utilisateurs et 15 items le tableau suivant montre les informations de chaque apprenant :

	$S_{t1}$	$S_{t2}$	$S_{t3}$	$S_{t4}$	$S_{t5}$	$t_{s1}$	$t_{s2}$	$t_{s3}$	$t_{s4}$	$t_{s5}$	$mc_x$
$a_1$	0.9	0.4	0.1	0.9	0.2	3	2	5	6	4	0.7
$a_2$	0.5	0.4	0.1	0.7	0.3	5	7	6	8	5	1.0
$a_3$	0.2	0.7	0.1	0.3	0.1	2	5	7	9	6	0.4
$a_4$	0.6	0.3	0.5	0.2	0.9	6	3	8	9	7	0.6
$a_5$	0.1	0.4	0.2	0.0	0.0	1	2	9	9	8	0.9
$a_6$	0.9	0.9	0.1	0.1	0.1	2	5	6	3	7	0.7
$a_7$	0.2	0.1	0.5	0.3	0.2	2	5	7	1	2	1.0
$a_8$	0.7	1.0	0.2	0.5	0.3	2	9	9	5	7	0.3
$a_9$	0.9	0.1	0.1	0.1	0.4	2	5	6	2	5	0.5
$a_{10}$	1.0	0.5	0.8	0.9	0.8	1	2	1	3	3	0.3

*Tableau 4.1. Informations des apprenants*

Avec  $S_{ti}$  = le score du test  $i$ ,  $t_{si}$  = le temps de passassions du test  $i$ ,  $mc_x$  : la capacité mémoire de l'apprenant.

Les évaluations qu'ont faites les apprenants sont montrées dans le tableau suivant :

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$	$a_{10}$
$t_1$	0.7	0.8	0.9	0.0	1.0	0.2	0.0	0.9	0.2	0.9
$t_2$	1.0	0.2	0.1	0.5	0.0	0.1	0.5	0.9	0.1	0.1
$t_3$	0.1	0.1	0.1	0.1	0.5	0.5	0.2	0.1	0.2	0.9
$t_4$	0.4	0.5	0.7	0.3	0.2	0.8	0.3	0.1	0.1	0.7
$t_5$	0.5	0.8	0.6	0.1	0.3	0.2	0.9	0.1	0.4	1.0
$t_6$	0.9	0.2	0.1	0.5	0.9	0.4	0.9	0.4	0.1	0.5
$t_7$	0.0	0.4	0.8	0.8	0.9	0.1	0.1	0.0	0.3	0.8
$t_8$	0.4	0.1	0.2	0.2	0.1	0.3	1.0	0.9	0.4	0.9
$t_9$	0.0	0.7	0.1	0.1	0.1	0.4	0.0	0.1	0.1	0.0
$t_{10}$	0.9	0.4	0.5	0.5	0.0	0.5	0.5	0.1	0.3	0.9
$t_{11}$	0.7	0.5	0.8	0.8	0.4	0.8	0.2	0.4	0.0	0.7
$t_{12}$	1.0	0.4	0.7	0.7	0.0	0.7	0.3	0.0	0.5	1.0
$t_{13}$	0.0	0.8	0.4	0.4	0.9	0.4	0.9	0.9	0.2	0.6
$t_{14}$	0.5	0.5	0.1	0.1	0.7	0.1	0.3	0.7	0.3	0.4
$t_{15}$	0.8	0.4	0.3	0.3	1.0	0.3	0.9	1.0	0.0	0.5

*Tableau 4.2. Les évaluations*



Un apprenant parmi les dix, soit a1, représente notre apprenant central, c'est-à-dire sur lequel se déroule le test. Notons que parmi les 15 items, 5 sont mal appréciés (note < 5). La liste des évaluations qu'a données a1 aux 15 items restants sur le plan réel est présentée dans le tableau suivant :

Item	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	15
évaluation	0.7	1.0	0.1	0.4	0.5	0.9	0.0	0.4	0.8	0.9	0.7	1.0	0.0	0.5	0.8

*Tableau 4.3. Les évaluations de l'apprenant a1*

### IV.3.1. Les résultats de l'expérimentation

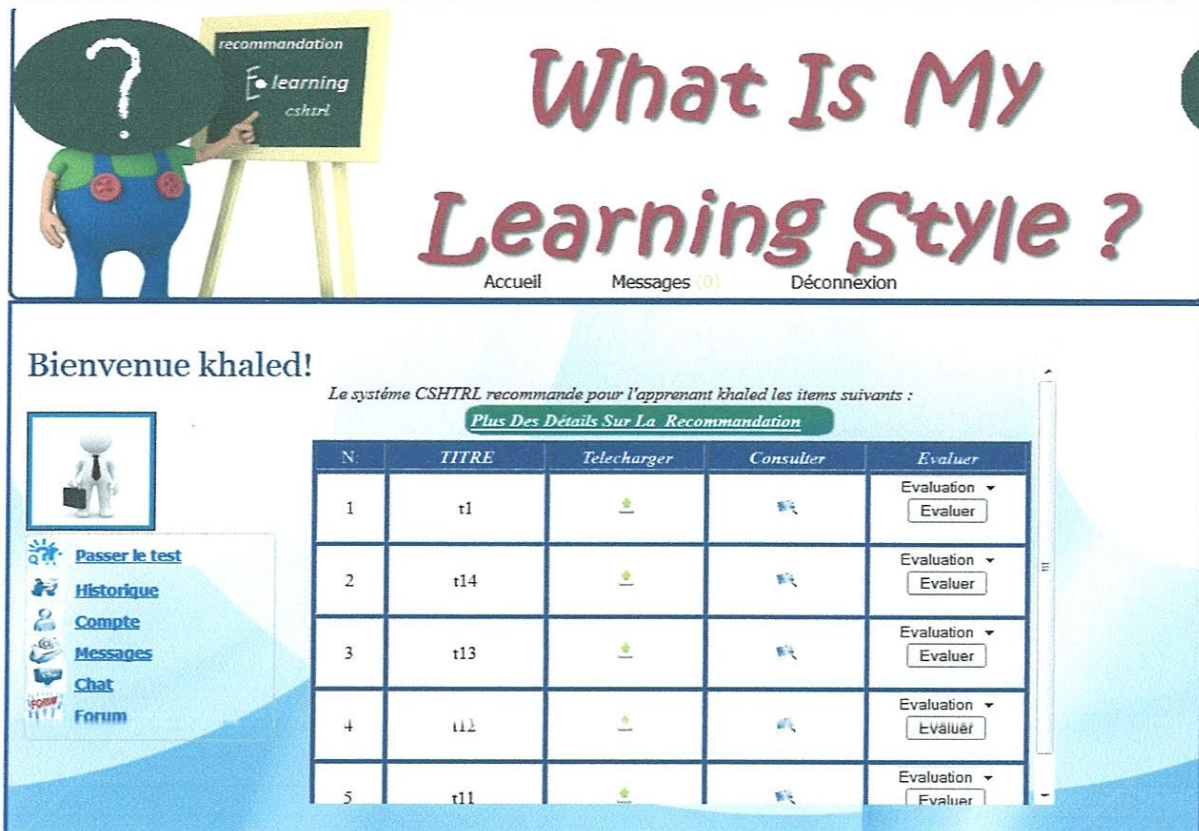
Les recommandations qui ont été générées pour a1 à partir de ces informations (évaluations qu'il a faites, scores des tests, capacité mémoire...), sont listées dans le tableau ci-dessous :

k	Les items recommandés
k=5	t <sub>1</sub> , t <sub>11</sub> , t <sub>4</sub> , t <sub>5</sub> , t <sub>12</sub>
k=10	t <sub>1</sub> , t <sub>11</sub> , t <sub>4</sub> , t <sub>5</sub> , t <sub>12</sub> , t <sub>10</sub> , t <sub>7</sub> , t <sub>13</sub> , t <sub>3</sub> , t <sub>9</sub>
K= 15	t <sub>1</sub> , t <sub>11</sub> , t <sub>4</sub> , t <sub>5</sub> , t <sub>12</sub> , t <sub>10</sub> , t <sub>7</sub> , t <sub>13</sub> , t <sub>3</sub> , t <sub>9</sub> , t <sub>8</sub> , t <sub>15</sub> , t <sub>6</sub> , t <sub>14</sub> , t <sub>2</sub>

*Tableau 4.4. Les items recommandés à l'apprenant a1*

La figure suivante montre le résultat de la recommandation pour k=5, et pour plus de détail sur la recommandation on clique sur « plus des détails sur la recommandation ».





*Figure 4.8 : Interface du résultat*

### IV.3.2. Posttest

On fait un post-test pour 5 apprenants, et chaque apprenant qui passer le test il faut rependu aux deux questions (donc si une repense juste alors la note est 5, si les deux repenses juste la note est 10, sinon la note est 0). Les résultats sont affichés dans le tableau suivant :

Apprenant(i)	1	2	3	4	5
Notes	5	10	10	5	10

*Tableau 4.5. Résultat posttest*

### IV.4. Conclusion

Dans ce chapitre nous avons présentés l'implémentation de notre travail, qui nous donne de bonne résultat si en compare avec les autres méthodes de recommandation (CSHTR) et prouvé que CSHTML produit une recommandation personnalisée de qualité.

A partir des résultats obtenus dans le post-test on peut dire que le système est efficace.

## Conclusion générale

Dans le cadre de ce mémoire, nous avons étudié principalement deux domaines (E-learning, et système de recommandation) afin de réaliser la nouvelle démarche de l'utilisation de systèmes de recommandation en e-learning. Ce système permet de filtrer les objets d'apprentissage selon les profils des apprenants.

Nos objectifs initiaux étaient d'utiliser les systèmes de recommandation dans le domaine du e-learning en appuyant sur la technique CSHTRL pour fournir aux étudiants de l'information pertinente.

L'expérimentation que nous avons réalisée prouvé que CSHTRL produit une recommandation personnalisée de qualité, mais elles ne peuvent être encore fixées définitivement.

Comme perspectives, nous espérons perfectionner les caractéristiques d'apprenants qui peuvent améliorer la qualité d'apprentissage et de recommandation.

Egalement, nous souhaitons enrichir notre technique de recommandation par des formules mathématiques qui peuvent améliorer notre système.

## Références et bibliographie

[Abbes, 1999] :Abbes, R;*Le filtrage des informations.Note de synthèse 1999*. 39 feuillets.  
Cote 1999 SYN DEA 01.Mars 1999.

Disponible à l'adresse:<http://enssibal.enssib.fr/bibliotheque/documents/dea/nsramzi.pdf>

[Baajour, 2006]:Baajour, H and Magoula, G (2006): *User modeling to support personalisation in the context of lifelong learning*. Titre: Life Long in London for All,  
disponible à l'adresse: [http:// www.lkl.ac.uk/research/14all](http://www.lkl.ac.uk/research/14all)

[Benhamdi, 2011]: S. Benhamdi, H. Seridi: Pedagogical content personalization, ITHET'1,  
vol(10), Turkey, 2011.

[Berrut, 2003] :Berrut, C. (2003) · *Filtrage collaboratif*. Chapitre 8, p255-283, F.Gaussier,  
M.H. Stefanini, Assistance intelligente à la recherche d'informations, Hermes-Lavoisier

[Chen et Mizoguchi, 1999]: Chen, W and Mizoguchi, R. (1999): *Communication Content Ontology for Learner Model Agent in Multi-agent Architecture*.

In Prof.AIED99 Workshop on Ontologies for Intelligent educational Systems. Disponible à  
l'adresse : <http://www.ei.sanken.osaka-u.ac.jp/aied99/a-papers/W-Chen.pdf>

[Chen et Sycara, 1998]: Chen, L. and Sycara, K. (1998): *Web Mate: A Personal Agent for Browsing and Searching*. Proceedings of the 2nd International Conference on Autonomous Agents and Multi Agent Systems, AGENTS '98, ACM, p.32-139.

[Dcmi, 2004]:Dcmi. (2004) *Dublin Core Metadata Initiative Usage Board, DCMI Metadata Terms*. Disponible à l'adresse: <http://dublincore.org/documents/dcmi-terms/>

[Dolog, 2004]:Dolog, P (2004): *Identifying Relevant Fragments of Learner Profile on the Semantic Web*. In The Workshop on Applications of Semantic Web technologies for

e-Learning SW-EL'04: Semantic Web for E-Learning at International semantic Web Conference: ISWC 2004 Hiroshima, Japan



**[Dologet Schafer, 2005]:**Dolog, P; Schäfer, M (2005): *A Framework for Browsing Manipulating and Maintaining Interoperable Learner Profiles*. In The UM 2005 Conference Edinburgh: Springer Verlag

**[Dumbill, 2002]:**Dumbill, E (2002): *XML Watch: Finding friends with XML and RDF: The Friend-of-a-Friend vocabulary can make it easier to manage online communities*.

**[Favier, 2001] :** Favier, L ; Morin, A (2001): *Filtrage statistique des documents textuels* , LIMSIC Université de Bourgogne, IRISA Institut de Recherche en Informatique et Systèmes Aléatoires, p 253-264.

**[Gallardo, 2005]:**Gallardo Lopez, M.L. (2005) : *Accès à l'Information par un Système de Filtrage Collaboratif Contrôlé*. Université Joseph Fourier,Grenoble I, Laboratoire CLIPS IMAG.

**[IMS LIP, 2001]:***IMS Learner Information Package Specification*, IMS Global Learning Consortium, 2001. Available at: <http://www.imsproject.org/profiles/index.html>

**[Kassem et al, 2004] :** Kassem, W.; Mounajed, A.; Saadoun, N. (2004) : *Etat de l'Art du E-Learning*. Projet du Management et NTIC. Université Pantheon-Assas, Paris II.

**[Lee, 2005]:** Lee, J.S; Jun, C.H; Lee, J et Kim, S. (2005): *Classification-based collaborative filtering using market basket data*. Department of Industrial Engineering,Pohang University of Science and Technology, San 31 Hyoja-dong, Pohang 790-784, South Korea, p700-704

**[LOM, 2002]:** Final Draft Standard for Learning Object Metadata, Approved draft,Document IEEE 1484.12.1-2002, 2002. 44p. Disponible à l'adresse: <http://ltsc.ieee.org/wg12/>

**[Maltz, 1995]:**Maltz, D.A; Ehrlich, K. (1995): *Pointing the way: active collaborative filtering*. Proceedings of CHI'95 , p 7-11.

**[Middleton, 2004]:** Middleton, S.E; Shadbolt, N.R.; De Roure, D. C (2004): *Ontological User Profiling in Recommender Systems*. University of Southampton, Dept., ACM,Inc., 1515, Broadway, New York, NY 10036 USA , ACM Transactions on Information Systems, Vol. 22, No. 1, January 2004, p 54-88.

[**Mobasher, 2006**]: Mobasher, B. (2006): *Data Mining for Personalization. In The Adaptive Web: Methods and Strategies of Web Personalization*. Brusilovsky, Kobsa, NejdI (eds.). Lecture Notes in Computer Science, Vol.4321. Springer-Verlag, Berlin Heidelberg, 2006.

[**Najjar et al, 2003**]: Najjar, J; Duval, E; Ternier, S; Neven, F.: *Towards Interoperable Learning Object Repositories: the ARIADNE Experience*. The IADIS International Conference WWW/Internet 2003, Algarve, Portugal, 5-8 Nov. 2003, Vol. I, pp. 219-226. Disponible à l'adresse : <http://www.cs.kuleuven.ac.be/~stefaan/papers/TowardsInteroperable.pdf>

[**Nouali, 2006**]: Nouali, O (2006): *Filtrage d'information*. Cours du Magister en Informatique, spécialité 'SIC' à l'université de Tlemcen.

[**Oubahssi, 2005**]: OUBAHSSI, L. (2005) : *Conception de plates-formes logicielles pour la formation à distance, présentant des propriétés d'adaptabilité à différentes catégories d'utilisateurs et d'interopérabilité avec d'autres environnements logiciels*. Thèse de doctorat de l'Université René Descartes – Paris V, Soutenue le 01 décembre 2005.

[**Ouraiba, 2008**]: Ouraiba E A, (2008): *Personnalisation de contenu en e-learning: Recommandation automatique des objets d'apprentissage aux apprenants en utilisant SVM*. Mémoire de magistère de l'université Abou Bekr Belkaid-Tlemcen, set enue en Avril.

[**PAPI, 2000**]: PAPI (2000): *Public and Private Information for Learners (PAPILearner)*. Draft Standard for Learning Technology. IEEE P1484.2/D7, 28/11/2000.

[**Rodríguez, 2002**]: Rodríguez-Artacho, M. (2002): *PALO Language Overview*(Draft). Disponible à l'adresse: <http://sensei.lsi.uned.es/palo/PALO-TR.pdf>

[**UCAID, 2002**]: UCAID (2002): *EduPerson Object Class Specification*. Disponible à l'adresse: <http://www.nmi-edit.org/eduPerson/internet2-mace-dir-eduPerson-200210.pdf>.

[**Vidal et al, 2004**] : Vidal, P ; Broisiiu, J ; Duval, E ; Ternier, S (2004): *Normalisation et Standardisation des Objets d'Apprentissage : L'Expérience ARIADNE*. Disponible sur : <http://e-miage.ups-tlse.fr/colloque/papiers/P-Vidal-ObjetApprentissage.pdf>