



République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**Université Abderrahmane Mira de Bejaia**  
Faculté des Sciences Exactes  
Département d'Informatique

ECOLE DOCTORALE RESEAUX ET SYSTEMES DISTRIBUES

## *Mémoire de Magistère*

**En Informatique**

**Option : Réseaux et Systèmes Distribués**

### *Thème*

---

## **Tagging Collaboratif et Filtrage de Tags à base du Profil Utilisateur**

---

Présenté par

**Mme DAHAK-KICHOU Saida**

Devant le jury composé de :

<b>Président</b>	TARI Abdelkamel	M.C.A	Université de Bejaia
<b>Rapporteur</b>	AMGHAR Youssef	Professeur	INSA Lyon France
<b>Examineur</b>	DRIAS Habiba	Professeur	USTHB Alger
<b>Examineur</b>	MAHDAOUI Latifa	M.C.A	USTHB Alger
<b>Invitée</b>	MELLAH Hakima	Chargée de recherche	CERIST Alger

**Promotion : 2008/2009**

# Remerciements

*Merci au bon Dieu, le tout Puissant.*

*Je tiens à remercier vivement mon directeur de thèse Mr Amghar Youssef, professeur à l'INSA, pour ses corrections, conseils et orientations. Merci d'avoir accepté de m'encadrer et me guider dans le monde si vaste de la recherche. Qu'il trouve en ce mémoire l'expression de mon profond respect.*

*Je remercie tout particulièrement Mme Mellah Hakjma, chargée de recherche au CERIST, codirecteur de thèse, de m'avoir orienté, corrigé mon travail, soutenu et encouragé. Merci pour sa disponibilité et sa gentillesse sans égale. Je ne la remercierai jamais assez, qu'elle trouve en ce mémoire l'expression de ma profonde gratitude et mon respect infini.*

*Mes vifs remerciements et respects s'adressent aux membres du jury Mr Tari Abdelkamel, maitre de conférence à l'université de Bejaia, Mme Drias Habiba, professeur à l'USTHB et Mme Mahdaoui Latifa, maitre de conférence à L'USTHB pour m'avoir fait l'honneur d'accepter de juger ce modeste travail.*

*Je tiens également à remercier Mr Badache Nadjib, professeur à l'USTHB et Directeur du CERIST, de m'avoir donné la chance d'entrer dans le monde de la recherche.*

*Mes sincères remerciements vont également à Mr Meziane Abdelkrim, chargé de recherche et responsable de la division systèmes d'information et systèmes multimédias pour ses encouragements et sa compréhension.*

*Je remercie infiniment Mr Dahak Fouad, maitre assistant à l'école supérieure d'informatique pour son aide très précieuse, ainsi que Mr Azouaou Faïçal, maitre de conférence pour ses orientations et conseils.*

*Mes chaleureux remerciements vont à mes très chers parents qui ne cessent de croire en moi et m'encourager. Merci pour toute ma famille, grands parents, oncles et tantes, cousins et cousines, et toute ma belle famille.*

*Mes chaleureux remerciements vont également à mes amis du CERIST Nour El Houda, Linda, Lamia, Fouzia, Safia, Amel et Leila pour leurs soutiens et encouragements. Merci à tous mes collègues de la division Imen, Badiaa, F.Zohra, Lamia, Djalila, Lydia, Salem, Kamal et Nouredine pour leurs encouragements.*

*Je remercie mes collègues de la promotion, Lamia, Sahar, Dalila, Faiza, Nabil, Madjid et Rafik pour la bonne ambiance vécue tout au long d'une année de cette belle expérience.*

*Merci à Mme Zaidi et tous les collègues du service formation pour leur disponibilité.*

*Je remercie mes amis Malika.B, Malika.A, Radia, Sekoura, Soraya et Fazia pour le soutien moral tant apprécié.*

*Fouad, merci encore une fois, je ne te remercierai jamais assez pour tout ce que tu fais pour moi, merci d'être à mes côtés dans les moments difficiles, sans toi je n'aurais jamais pu tenir jusqu'à la fin.*

*Tout simplement Merci !*

*A mes très chers parents Rahima et Abdelmadjid*

*A mon mari Fouad*

*A mes deux adorables petits anges Anis et Yacine*

*A ma chère sœur Samia*

*A mes frères Mohamed et Azedine*

*A mes beaux parents Said et Wardia*

*A jida Adidi, jeddi Omar et ami Mustapha*

*A toute ma famille et belle famille*

## RESUME

Le 'Tagging collaboratif' ne cesse de gagner une popularité sur le web 2.0, cette nouvelle génération du web qui fait de l'utilisateur un lecteur-rédacteur. Le 'Tagging' est un des moyens permettant à l'utilisateur de s'exprimer librement via des ajouts d'étiquettes appelées 'Tags' à des ressources partagées. L'un des problèmes rencontrés dans les systèmes du Tagging actuels est la définition des tags les plus appropriés pour une ressource. Les tags sont généralement classés par ordre de popularité tel que del-icio-us. Or la popularité du tag ne reflète pas toujours son importance et sa représentativité vis-à-vis de la ressource à laquelle il est associé. Partant des hypothèses qu'un même tag pour une ressource peut prendre des significations différentes selon les utilisateurs, et un tag issu d'un utilisateur connaisseur serait plus important qu'un tag issu d'un utilisateur novice, nous proposons une approche de pondération des tags d'une ressource à base du profil utilisateur. Pour ceci nous définissons un modèle utilisateur permettant son intégration dans le calcul du poids d'un tag ainsi qu'une formule de calcul de ce dernier à base de trois facteurs concernant l'utilisateur à savoir, le degré de rapprochement entre ses centres d'intérêts et le domaine de la ressource, son expertise et son estimation personnelle vis-à-vis des tags qu'il associe à la ressource. Un descripteur de ressource contenant les meilleurs tags est ainsi créé.

### *Mots-clés*

Annotation, Tagging collaboratif, profil utilisateur, recherche d'informations.

## ABSTRACT

The 'Collaborative Tagging' is gaining popularity on Web 2.0, this new generation of Web which makes user reader/writer. The 'Tagging' is a means for users to express themselves freely through additions of label called 'Tags' to shared resources. One of the problems encountered in current tagging systems is to define the most appropriate tag for a resource. Tags are typically listed in order of popularity, as del-icio-us. But the popularity of the tag does not always reflect its importance and representativeness for the resource to which it is associated. Starting from the assumptions that the same tag for a resource can take different meanings for different users, and a tag from a knowledgeable user would be more important than a tag from a novice user, we propose an approach for weighting resource tags based on user profile. For this we define a user model for its integration in calculating the weight of a tag and a formula for calculating it based on three factors namely the user, the degree of approximation between its centers interest and the field of resource, expertise and personal assessment for tags associated to the resource. A resource descriptor containing the best tags is created.

### *Keywords*

Annotation, Collaborative Tagging, user profile, information retrieval.

## ملخص

اكتسب التوسيم التعاوني شعبية كبيرة في الوب 2.0 ، النسخة الجديدة من الشبكة العنكبوتية التي جعلت المستخدم قارئ و محرر في آن واحد. التوسيم أحد الوسائل التي تسمح للمستخدم أن يعبر عن رأيه بكل حرية و ذلك بإضافة علامات بشكل كلمات تسمى 'Tag' لموارد مشتركة على الشبكة. من بين نقائص أنظمة التوسيم التعاوني الحالية التعريف بالعلامات المناسبة أكثر لمورد ما. هذه العلامات مرتبة حسب شعبيتها (عدد مرات ذكرها) مثل النظام Delicious. لكن شعبية العلامة لا تعكس دائما أهميتها و تمثيلها بالنسبة للمورد المنسوبة إليه. بافتراض أنه يمكن لعلامة ما أن تحمل معاني مختلفة حسب المستخدمين، و علامة مقترحة من مستخدم خبير أكثر أهمية من علامة مقترحة من مستخدم مبتدأ، نقترح نظاما لترجيح العلامات مستندا على مواصفات المستخدم. لذلك نقترح نموذجا لمواصفات المستخدم، يسمح هذا النموذج بإدخال المواصفات في عملية ترجيح العلامات. و صيغة لحساب وزن كل علامة مستندا على ثلاثة عوامل: درجة التقارب بين المورد و مراكز اهتمام المستخدم، خبرة المستخدم و تقييمه الشخصي لعلاماته. بذلك نكون مجموعة أحسن العلامات.

**الكلمات المفتاحية:** الملاحظات، التوسيم التعاوني، مواصفات المستخدم، استرجاع المعلومات.

# Table des Matières

## Introduction Générale

<b>Introduction</b> .....	<b>1</b>
<b>Contexte du travail et problématique</b> .....	<b>1</b>
<b>Contribution</b> .....	<b>2</b>
<b>Organisation du mémoire</b> .....	<b>3</b>

## Première Partie:Etat de l'Art

### Chapitre I : Les Annotations

<b>I.1. Introduction</b> .....	<b>6</b>
<b>I.2. Définitions</b> .....	<b>6</b>
<b>I.3. Structure de l'annotation</b> .....	<b>7</b>
I.3.1 L'objet Annotation .....	7
I.3.2. L'activité Annotation .....	11
<b>I.4. Sémantique de l'annotation (objectifs)</b> .....	<b>12</b>
I.4.1. Les objectifs selon C. Marshall [Marshall, 98] .....	12
I.4.2. Les objectifs selon J. Virbel [Veron, 97] .....	13
I.4.3. Les objectifs selon [Mille, 05].....	14
<b>I.5. Catégorisation des annotations</b> .....	<b>14</b>
I.5.1. Les catégories de l'objet annotation.....	15
I.5.1.1. L'annotation cognitive.....	15
I.5.1.2. L'annotation computationnelle.....	15
I.5.1.3. L'annotation sémantique .....	15
I.5.2. Les catégories de l'activité annotation .....	15
I.5.2.1.L'annotation manuelle.....	15
I.5.2.2. L'annotation semi-automatique.....	16
I.5.2.3. L'annotation automatique.....	16
<b>I.6. Les outils d'annotations</b> .....	<b>16</b>
<b>I.7. L'annotation sémantique</b> .....	<b>18</b>
I.7.1. Définitions.....	18
I.7.2. Les langages d'annotation sémantique.....	19
<b>I.8. Conclusion</b> .....	<b>19</b>

## Chapitre II : Le Tagging Collaboratif

<b>II.1. Introduction</b> .....	<b>21</b>
<b>II.2. Définitions</b> .....	<b>21</b>
II.2.1. Le Tagging collaboratif.....	21
II.2.2. Le Tag .....	22
II.2.3. La Folksonomie .....	22
<b>II.3. Structure d'une action du Tagging collaboratif</b> .....	<b>24</b>
II.3.1. Structure tripartite de base .....	24
II.3.2. Structure tripartite avec liens inter-ressources et inter-utilisateurs.....	25
II.3.3. Structure quadripartite .....	26
<b>II.4. Propriétés d'un système du Tagging collaboratif</b> .....	<b>27</b>
<b>II.5. Tagging vs Annotation</b> .....	<b>28</b>
<b>II.6. Etude des systèmes du Tagging collaboratif</b> .....	<b>28</b>
II.6.1. Etude de la dynamique des systèmes du Tagging.....	29
II.6.2. Travaux sur la proposition de modèles et d'algorithmes de suggestion de tags et d'utilisateurs .....	30
II.6.3. Découverte de communauté.....	31
<b>II.7. Tagging collaboratif et recherche d'information</b> .....	<b>31</b>
<b>II.8. Rapprocher les ontologies et les folksonomies</b> .....	<b>33</b>
II.8.1. Les approches d'extraction de liens sémantiques entre tags.....	33
II.8.1.1. Analyse des réseaux sociaux appliquée aux folksonomies .....	33
II.8.1.2. Analyse de la dynamique des folksonomies .....	33
II.8.1.3. Clustering .....	34
II.8.2. Les approches basées sur les ontologies .....	34
II.8.2.1. Guider le Tagging à l'aide d'ontologies.....	34
II.8.2.2. Construire une ontologie de folksonomies.....	34
II.8.3. Exemples d'ontologies informatiques pour le Tagging .....	34
<b>II.9. Les limites des systèmes du Tagging</b> .....	<b>36</b>
<b>II.10. Conclusion</b> .....	<b>37</b>

## Chapitre III : Le Profil Utilisateur

<b>III.1. Introduction</b> .....	<b>38</b>
<b>III.2. Définition</b> .....	<b>38</b>
<b>III.3. Modélisation du Profil</b> .....	<b>39</b>
<b>III.4. Représentation du Profil</b> .....	<b>40</b>
III.4.1. Représentation ensembliste ou vectorielle.....	40
III.4.2. Représentation sémantique .....	40
III.4.3. Représentation connexionniste .....	40

III.4.4. Représentation multidimensionnelle.....	40
III.4.5. Représentation hiérarchique.....	40
<b>III.5. Dimensions d'un Profil utilisateur .....</b>	<b>41</b>
<b>III.6. Acquisition du Profil .....</b>	<b>42</b>
III.6.1. Approche simpliste : .....	43
III.6.2. Approche dynamique : .....	43
III.6.3. Approche par apprentissage : .....	43
<b>III.7. Modèle conceptuel du profil .....</b>	<b>43</b>
<b>III.8. Profil utilisateur dans le Web 2.0 .....</b>	<b>44</b>
III.8.1. Recherche à base de tags et profil utilisateur .....	45
III.8.2. Créer et enrichir le profil utilisateur en se basant sur les tags .....	45
<b>III.9. Conclusion .....</b>	<b>47</b>

## Deuxième Partie

### Une Approche de Filtrage de Tags à base du Profil utilisateur

#### Chapitre IV : Présentation de l'approche

<b>IV.1. Introduction.....</b>	<b>49</b>
<b>IV.2. Motivations .....</b>	<b>50</b>
<b>IV.3 Principe général .....</b>	<b>50</b>
<b>IV.4 Présentation de l'approche .....</b>	<b>51</b>
IV.4.1. Le modèle du profil utilisateur .....	53
IV.4.1.1. Représentation du profil .....	53
IV.4.1.1.1. La dimension personnelle.....	52
IV.4.1.1.2. La dimension centres d'intérêts.....	52
IV.4.1.1.3. La dimension expertise.....	52
IV.4.1.2. Construction du profil.....	55
IV.4.1.2.1. Construction de la dimension centres d'intérêts.....	54
- L'approche hybride.....	54
- L'algorithme Add-A-Tag adapté à notre approche hybride.....	56
IV.4.1.2.2. Construction de la dimension expertise.....	56
- Choix de l'ontologie.....	57
- Profondeurs des termes dans WordNet.....	58
IV.4.2. Pondération des tags à base du profil utilisateur .....	60
IV.4.2.1. Etude des variations de la formule de pondération.....	63
IV.4.3. Classement des tags et construction des descripteurs (filtrage).....	65
IV.5 Conclusion.....	66

#### Chapitre V : Tests et Evaluations



<b>V.1. Introduction .....</b>	<b>67</b>
<b>V.2. Collection de test.....</b>	<b>67</b>
<b>V.3. Démarche d'évaluation .....</b>	<b>67</b>
V.3.1. Phase de préparation de la collection.....	68
V.3.1.1. Elimination des tags ne figurant pas dans WordNet.....	68
V.3.1.2. Récupération des profondeurs des tags à partir de WordNet.....	68
V.3.1.3. Indexation du contenu textuel des pages web.....	69
V.3.1.3. Elimination des mots-clés de l'index ne figurant pas dans WordNet.....	71
V.3.2. Evaluation en utilisant un système de recherche d'information (SRI).....	71
V.3.2.1. Implémentation du moteur de recherche.....	71
V.3.2.2. Construction du vecteur idéal .....	72
V.3.2.3. Construction du vecteur popularité .....	73
V.3.2.4. Construction du vecteur poids .....	74
V.3.3. Comparaison des résultats .....	75
<b>V.4. Architecture du système d'évaluation .....</b>	<b>76</b>
<b>V.5. Résultats et discussion.....</b>	<b>79</b>
<b>V.6. Conclusion.....</b>	<b>81</b>

### **Conclusion Générale**

<b>Synthèse.....</b>	<b>82</b>
<b>Résumé de la contribution.....</b>	<b>82</b>
<b>Perspectives.....</b>	<b>83</b>

### **Bibliographie**

### **Annexes**