



الجمهورية الجزائرية الديمقراطية الشعبية



République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة قاصدي مرباح - ورقلة-

Université Kasdi Merbah -Ouargla

Faculté des Sciences et de la Technologie et Sciences de la Matière

Département des Mathématiques et d'Informatique

MEMOIRE DE FIN D'ETUDES

En vue de l'obtention du

Diplôme d'ingénieur d'état en Informatique

Option: Informatique Industrielle

Approche du tagging collaboratif pour la navigation sociale et la structuration de l'information sur Internet

Présenté par :

➤ BABELHADJ Yamina

➤ SALHI Amina

Dirigé par :

M. HEROUZ Abd Elhakim

Session : Juin 2010

Remerciements

Tout d'abord, nous tenons à remercier **DIEU**, notre créateur pour nous avoir donné la force à accomplir ce travail.

Nous remercions vivement Monsieur : **Herrouz Abd Elhakim** notre encadreur pour son aide, sa compréhension, ses conseils et critiques pertinentes.

Nous remercions également toutes les personnes qui nous ont aidés, de près ou de loin pour la réalisation de ce travail.

Nous tenons à remercier les membres du jury pour l'honneur qu'ils nous ont fait en acceptant de juger notre travail.

Nous adressons aussi nos remerciements à tous nos enseignants qui ont veillé sur notre formation.

Amina et Yamina ... 

TABLE DES MATIERES

Remerciements	I
Table des matières	II
Table des figures.....	VIII
Liste des tableaux.....	XI
Résumé.....	01
Introduction générale.....	03

Chapitre 1 : Concept général de web2.0

Introduction.....	07
1.1 Qu'est ce que web 2.0.....	08
1.1.1 Définition web 2.0.....	08
1.1.2 Naissance du concept	09
1.1.3 Les principes de web 2.0.....	10
1.1.4 Les évolutions par rapport au web 1.0 : le rôle de l'internaute	12
1.1.5 Web 2.0 concepts pratiques et technologiques	14
1.1.6 Glossaire du Web 2.0	16
1.2 L'intelligence collective et le web 2.0	19
1.3 Les Nouveautés Du Web 2.0.....	21
1.4 Les limites du web 2.0	23
1.5 Pourquoi le Web 2.0 marche-t-il	25

1.6	L'évolution du Web : la défragmentation de l'Infosphère	27
1.7	Les web sémantiques	29
	Conclusion.....	31

Chapitre 2 : Le tagging collaboratif

	Introduction.....	34
2.1	tagging collaboratif	35
2.2	les tags	37
2.2.1	La recherche par tags	38
2.2.2	Partageons les tags	38
2.2.3	Tags structurables	41
2.2.4	Le concept de noyau : tagging.....	42
2.2.5	Pourquoi les utilisateurs étiquettent-ils.....	46
2.3	Folksonomie	48
2.3.1	Définition et présentation de la folksonomie.....	48
2.3.2	Folksonomy comme données qui sont émergentes de l'information partagée	49
2.3.2	Traitements sur les folksonomies	50
2.4	Tagging et Taxonomie	51
2.5	Le tagging et le web sémantique	53
	Conclusion.....	55

Chapitre 3 : Présentation des systèmes de tagging existants

	Introduction.....	58
3.1	Delicious.....	60

3.1.1	A quoi sert delicious ?.....	60
3.1.2	dynamique Delicious	61
3.1.2.1	Les données	63
3.1.2.2	Activité de l'utilisateur et la quantité tag	63
3.1.2.3	Types de tags	67
3.1.3	Ses points forts	69
3.1.4	Ses points faibles	70
3.1.5	Utiliser Delicous sans s'y inscrire	70
3.1.6	Création d'un compte delicious	73
3.1.7	Ajout les favoris	83
3.1.7.1	Manuellement	83
3.1.7.2	Utiliser les boutons de bookmark.....	85
3.1.7.3	Importer vos favoris.....	86
3.1.8	Organisation des tags	88
3.1.9	Options d'affichage des tags	90
3.1.10	Optimiser l'utilisation de Delicious dans un travail d'équipe en association ou ONG	91
3.2	Flickr	93
3.2.1	Utilisation de tag	96
3.2.2	Créer un compte sur flickr	98
3.2.3	Importer les photos dans ce compte de flickr	101
3.3	YouTube.....	103

3.3.1	Création d'un compte sur YouTube.....	104
3.3.2	Chargement de votre vidéo sur votre compte.....	106
3.3.3	L'utilisation des tags par Youtube	108

Chapitre 4 : La navigation sur le Web

	Introduction.....	112
4.1	Qu'est-ce que la "navigation" ?	113
4.2	L'hypertexte.....	113
4.2.2.1	Hypertexte : nœud et liens.....	115
4.2.2.2	Propriétés de l'hypertexte	116
4.2.2.3	Inconvénients de l'hypertexte	117
4.2.2.4	Lien hypertexte (ou <i>hyperlien</i>).....	118
4.3	Hypermédia.....	118
4.4	Hyperdocument.....	119
4.5	Navigation : espace de conception	120
4.5.1	Navigation spatiale.....	120
4.5.2	Navigation sémantique	121
4.5.3	Navigation sociale	122
4.6	Navigation par tags	124
4.6.1	Les tags, un outil de description et de navigation.....	124
4.6.2	Utilisation des tags.....	125

4.6.3	L'intérêt des nuages de tags pour simplifier la navigation.....	125
4.6.4	Résultats en termes de référencement.....	126
4.7	Navigation et Social Tagging des systèmes.....	127
4.8	Systèmes de Tagging	128

Chapitre 5 : Réalisation

5.1	Les motivations des langages	132
5.1.1	Python	132
5.1.2	Django.....	134
5.2	L'installation l'environnement de travail.....	135
5.2.1	Installer python sous Windows	135
5.2.2	Installer django sous Windows	135
5.3	Description de l'application « Djanogo Bookmarks ».....	136
5.3.1	L'interface de l'application	136
5.3.2	Sauvegarder les bookmarks.....	139
5.3.3	La recherche de bookmarks.....	139
5.3.4	Bookmarks d'utilisateur	140
	Conclusion générale.....	143
	Bibliographie.....	145
	Abréviations.....	151

TABLE DES FIGURES

N°Figure	Titre	Page
Figure 1.1	Power Law of Participation, par Ross Mayfield. CC.....	14
Figure 1.2	Le cycle de vie de l'information.....	27
Figure 1.3	Cartographie de l'évolution du Web.....	28
Figure 1.4	Web 1.0 et 2.0 comparés au Web Sémantique.....	29
Figure 2.1	Les principaux concepts.....	35
Figure 2.2	Exemple de ressource vidéo employée par les enseignants.....	41
Figure 2.3	L'Amazone permet aux utilisateurs de tags un produit et voir comment d'autres ont tagging le même produit.....	47
Figure 3.1	Le nombre de tags dans la liste de tag de chaque utilisateur, en par ordre décroissante.....	64
Figure 3.2	Deux utilisateurs extrêmes la croissance tag.....	65
Figure 3.3	Taux de croissance des trois balises sélectionné pour deux utilisateurs	66
Figure 3.4	Comme pour une balise dans un signet (horizontal) augmente, son rang (fréquence) dans la liste des tags (vertical) diminue.....	69
Figure 3.5	Moteur de recherche de delicious.....	70
Figure 3.6	Le résultat de la recherche pour le mot « microcredit ».....	71
Figure 3.7	Un partie de résultat de recherche.....	72
Figure 3.8	L'interface de site delicious.....	73
Figure 3.9	Bouton « Join Now » pour inscrire.....	73
Figure 3.10	Page de loging de yahoo.....	74
Figure 3.11	page de confirmation de l'inscription de delicious.....	74
Figure 3.12	étape d'installation des bookmarklets.....	75
Figure 3.14	démarrer la téléchargement de fichier.....	76

Figure 3.15	exécuter le fichier.....	76
Figure3.16	démarrer l'installation.....	76
Figure 3.17	terminer l'installation.....	77
Figure 3.18	demande l'installation de bookmark.....	77
Figure 3.19	demande l'autorisation.....	77
Figure 3.20	le module complémentaire.....	78
Figure 3.21	l'installation est en cours.....	78
Figure 3.22	redémarrer firefox pour terminer l'installation.....	79
Figure 3.23	terminé l'installation.....	79
Figure 3.24	les outils de delicious.....	79
Figure 3.25	page d'utilisateur sur delicious.....	80
Figure 3.26	les 3 sections de navigations sur delicious.....	81
Figure 3.27	moteur de recherche sur delicious.....	81
Figure 3.28	page d'accueil de delicious.....	82
Figure 3.29	page flickr.....	84
Figure 3.30	page save a bookmark.....	85
Figure 3.31	page import bookmark.....	85
Figure 3.32	l'importe est successif.....	86
Figure 3.33	Exemple de bookmark importé.....	87
Figure 3.34	Distribution (densité) du numéro des tags distincts par utilisateur.....	90
Figure 3.35	Page d'accueil de Flickr.....	98
Figure 3.36	Personnalisation du compte Flickr.....	99
Figure 3.37	Page d'accueil personnalisée de lina-photo.....	99
Figure 3.38	Page de parcourir le photo.....	101
Figure 3.39	Les photos importer.....	101
Figure 3.43	Page d'inscription de youtube.....	104

Figure 3.44	Page de logging de youtube.....	105
Figure 3.45	Page envoyer vidéo.....	106
Figure 3.46	Page de réussite d'envoyer vidéo.....	107
Figure 4.1	Les trois modèles de navigation.....	120
Figure 4.2	Les caractéristiques de la navigation.....	121
Figure 5.1	La page principale de site Django Bookmarks.....	136
Figure 5.2	Interface de l'accès au système et d'indentification de l'utilisateur.....	137
Figure 5.3	Interface d'inscription du système.....	137
Figure 5.4	Interface due succès du d'enregistrement.....	138
Figure 5.5	Interface utilisateur.....	138
Figure 5.6	Interface de Sauvegarder les bookmarks.....	139
Figure 5.7	Interface de la recherche.....	139
Figure 5.8	la fenêtre d'affichage des bookmarks d'un utilisateur.....	140
Figure 5.9	la fenêtre de la modification d'un Tag.....	140
Figure 5.10	la fenêtre de correspondante au mot web 2.0.....	141
Figure 5.11	nuage de Tag de l'application.....	141

LISTE DES TABLEAUX

N°Tableau	Titre	Page
Tableau 1.1	Du Web 1 vers le Web 2.....	12
Tableau 3.2	Corrélation d'utilisation de Flickr.....	96
Tableau 4.1	résultats de l'utilisation des tags sur deux sites.....	127

RESUME :

L'utilisation des tags offre une nouvelle façon de naviguer sur la Toile. Les systèmes de tagging sont étroitement liés à la mouvance Web 2.0 ou Web social. Les internautes étant les créateurs des contenus non seulement textuels mais aussi multimédia. Il fallait trouver un moyen de leur faire décrire les contenus publiés. Sans catégorisation imposée, cette approche accélère le processus de mise en ligne et la liberté de catégorisation de l'internaute. Ceci permet donc d'éviter un lourd travail éditorial.

Ce mémoire propose une étude des systèmes de tagging et il présente un système de navigation par les tags afin de monter les avantages de cette technique.

MOTS-CLES :

Web 2.0, navigation, tags, système de tagging, social tagging, nuages de tags

المخلص:

استخدام الوسم يوفر طريقة جديدة لتصفح الانترنت. هذه الطريقة ترتبط ارتباطا بالويب 2.0 و الشبكة الاجتماعية. مستخدمي الانترنت لا يبدعون في وضع المحتويات النصية فحسب ، بل أيضا الوسائط المتعددة. ذلك لايجاد وسيلة للسماح لهم في وصف المحتويات المنشورة هذا النهج يسرع عملية النشر على الانترنت وحرية تصنيف المحتويات من قبل المستخدم. وبالتالي تجنب اي عبء في عملية التحرير.

وتقترح هذه المذكرة دراسة نظم *tagging* ، وتعرض نظام لتصفح الانترنت بالبطاقات لتجميع مزايا هذه التقنية.

المفاتيح :

الويب 2.0 ، والملاحة ، الوسوم ، أنظمة الوسم ، المفضلات الاجتماعية ، غيوم الوسم

Resume:

The use of the tags offers a new way of navigate on the Web. The tagging systems are closely linked to the Web 2.0 movement and social Web. Customers are the creators of content not only textual but also multimedia. He had to find a way to let them describe the content published. Without imposed categorization, this approach accelerates the process of online freedom and categorization of the user. This solution avoids a heavy editorial work.

This memory proposes a study of the systems of tagging and it presents a navigation system by the tags in order to assemble the advantages of this technique.

The key words:

Web 2.0, navigation, tags, tagging system, social tagging, tag clouds

Introduction Générale

Introduction générale

Le Web est un espace informationnel ouvert, dynamique, distribué, hétérogène et non modéré. On constate que sur cet hypermédia, des sites apparaissent et disparaissent, des contenus sont modifiés et il devient impossible de maîtriser l'organisation. Par contre, Internet a rendu possible des situations dans lesquelles l'utilisateur construit lui-même son cheminement individualisé. On parle ainsi d'activité de navigation pour s'approprier un contenu ou se repérer et choisir des sources d'information. Cependant, si Internet offre un potentiel énorme de par le haut degré d'interactivité et de connectivité qu'il permet, n'en possède pas moins certaines limites relatives à la tâche de navigation.

En effet, le processus de navigation sur Internet se trouve confronté à de nombreux problèmes. D'une part, la nature de l'environnement lui-même impose ses inconvénients : Internet est un réseau à l'échelle mondiale, en constante évolution et non structuré. D'autre part, les internautes montrent généralement des difficultés à construire un schéma mental de navigation. Les systèmes d'aide à la navigation proposés par les différents logiciels traditionnels ne répondent pas aux besoins des usagers et contribuent parfois, paradoxalement, à rendre le processus plus confus. L'analyse de ces problèmes mène à la conclusion que le développement de nouveaux outils d'assistance à la navigation s'impose.

Actuellement, la navigation en pied de page devient de plus en plus populaire auprès des sites ayant une grande diversité de sujets à aborder. En général, il existe plusieurs modes de présentation de cette navigation, entre autres : affichée en pleine page sur la page d'accueil, et/ou en pied de page sur les pages internes ou en plus petit dans la colonne de droite ou de gauche.

Cette forme de navigation présente une liste de liens affichés à la suite les uns des autres séparés d'un espace régulier. Ces liens sont des mots-clés ou des intitulés d'articles et ils sont appelés tags. Un tag porte la notion de description informelle d'un contenu. Un ensemble de tags définit un contenu. Pour retrouver le contenu, il suffit de cliquer sur un de ces tags. La liste des tags est triée par ordre alphabétique, par ordre décroissant de création, ou autre.

Cette navigation présente une autre dimension : les libellés sont écrits en plus ou moins gros. Plus un libellé est gros, plus il est utilisé dans le site. Des variantes de ce concept existent. Par exemple, plus un tag est gros, plus il est populaire auprès des internautes.

Cette façon de présenter la navigation est intimement liée à la mouvance Web 2.0 ou Web social. Les internautes étant les créateurs des contenus non seulement textuels mais aussi multimédia. Il fallait trouver un moyen de leur faire décrire les contenus publiés. Sans catégorisation imposée, cette navigation accélère le processus de mise en ligne et la liberté de catégorisation de l'utilisateur. Ceci permet d'éviter un lourd travail éditorial. La catégorisation d'un site s'appelle la taxonomie des contenus. Les tags étant générés par les internautes dans une langue beaucoup plus informelle. C'est pourquoi, le web 2.0 a vu apparaître le terme de Folksonomie ou la catégorisation du peuple.

Ce travail propose une étude des systèmes de tagging et il présente un système de navigation par les tags afin de montrer les avantages de cette approche.

Ce document est organisé en cinq chapitres. Nous consacrerons un premier chapitre pour présenter le Web 2.0 ainsi que les technologies correspondantes.

Le second chapitre est consacré à l'étude des systèmes de tagging.

Le troisième chapitre portera sur la présentation des systèmes de tagging existants.

Dans le quatrième chapitre, nous donnons les définitions de certains termes les plus fréquemment utilisés dans ce mémoire : hypertexte, hypermédia, navigation et orientation. La notion de navigation hypertextuelle retiendra tout particulièrement notre attention. En fin de ce chapitre, nous aborderons l'approche de la navigation par les tags sur Internet.

Le cinquième chapitre, quant à lui, est consacré à la réalisation de notre système de tagging.

Enfin, en guise de conclusions, nous indiquerons quelques remarques sur ce travail en concluant par un bilan et nous exposerons les perspectives pour de futurs travaux.

Chapitre I

Généralités sur le Web2.0

Introduction :

Internet est sans aucun doute une des innovations majeures de l'histoire moderne. Comme pour toutes les grandes découvertes (par exemple l'électricité ou l'automobile), l'impact d'Internet sur notre société est d'ores et déjà considérable et les modes de vie de nos contemporains s'en trouvent profondément modifiés. Résultat de la fusion entre informatique et télécommunications, Internet constitue une vraie révolution technologique en permettant entre autres, un accès à la connaissance universelle, des échanges instantanés de données, une communication directe et permanente. Internet redéfinit des notions aussi fondamentales que celles du temps, de l'espace ou encore de la matérialisation. Parmi les inventions qui ont marquées l'Histoire, Internet est toutefois un cas unique de par la rapidité de sa *diffusion* (inventé il y a moins de quinze ans, Internet est aujourd'hui quotidiennement utilisé par plus d'un milliard de personnes à travers le monde) et de par la nature de son développement (des milliers de nouvelles applications naissent, en appelant d'autres toujours plus inattendues). A l'heure où l'on parle du *web 2.0* (internet de 2ème génération).

Plusieurs définitions ont été avancées pour expliquer ce que signifie le Web 2.0 mais ce qu'on l'on remarque quand on cherche une définition sur Internet c'est que tout le monde ne s'accorde pas sur le terme Web 2.0 et que ce concept ne fait pas l'unanimité.

Nous allons étudier ce phénomène en tentant de trouver une bonne définition, nous verrons également les technologies qui vont participer à cette « nouvelle version du web » et ce qui va changer pour les utilisateurs.

1.1. Qu'est ce que web 2.0 :



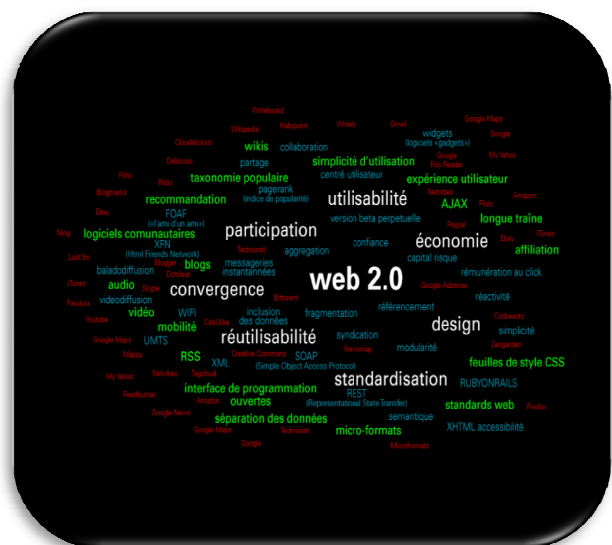
Selon le moteur de recherche Google, le terme est cité plus de 9,5 millions de fois en septembre 2005, plus de 573 millions en octobre 2007. Concept apparu en 2004, le web 2.0 a aujourd'hui dépassé le stade du simple buzz ("technique marketing consistant à faire du bruit sur un nouveau produit ou une offre"). Pour autant, un nombre non négligeable de définitions liées au web 2.0 restent floues ou

évoluent encore. La définition du mot lui-même, "web 2.0", proposée par Wikipédia depuis septembre 2005 est aujourd'hui modifiée plusieurs fois par semaine. L'expression « **Web 2.0** » désigne l'ensemble des technologies et des usages du *World Wide Web* qui ont suivi la forme initiale du web¹, en particulier les interfaces permettant aux internautes ayant peu de connaissances techniques de s'approprier les nouvelles fonctionnalités du web et ainsi d'interagir de façon simple à la fois avec le contenu et la structure des pages et aussi entre eux, créant ainsi notamment le Web social.

Cette expression utilisée par Dale Dougherty en 2003, diffusée par Tim O'Reilly en 2004 et consolidée en 2005 avec le position paper « What Is Web 2.0 » s'est imposée à partir de 2007.

1.1.1 Définition web 2.0:

Comment définir de façon précise ce que recouvre cette notion parfois floue ? Depuis plusieurs mois, le blog collectif "Dessine moi le Web 2.0 " propose à des universitaires, informaticiens, philosophes, entrepreneurs, de donner leur propre vision du phénomène et échanger idées et points de vue. La diversité et la richesse des



Contributions illustrent la difficulté à donner une définition exhaustive de ce concept à géométrie variable, mêlant à la fois des aspects techniques et sociaux.

Une vision réductrice du Web 2.0 consisterait à n'y voir que la mutualisation de contenus produits et mis en ligne par les utilisateurs eux-mêmes, ou à résumer l'engouement actuel à un effet de mode passager. Depuis la première conférence "Web 2.0 " en 2004, le concept s'est concrétisé à travers de nombreuses expérimentations ou le développement de nombreux services largement utilisés. Nous tenterons dans ce document de broser un panorama des enjeux, usages et technologies liés à cette notion aussi encensée que controversée.

Plusieurs auteurs s'accordent sur l'idée selon laquelle le web 2.0 ne constitue pas une rupture technologique ni ... sociale. Il s'agirait plutôt d'un retour aux sources, d'une renaissance du web tel qu'il a été conçu à l'origine, les utilisateurs s'étant entre temps appropriés les modes de création et de diffusion. Pour Paul Graham⁵, pionnier de l'Internet et fondateur de la société Viaweb en 1995, *"le web 2.0 est le web comme il a été conçu pour être utilisé. Les tendances que nous distinguons sont simplement la nature inhérente du web qui émerge des mauvaises pratiques qui lui ont été imposées pendant la bulle Internet"*.

Le web créé au début des années 90 par Tim Berners Lee comportait déjà en germe bien des éléments du web 2.0 : mutualisation des connaissances, échange, travail collaboratif, interactivité, "intelligence collective " ... Ce n'est sans doute pas un hasard si Dale Dougherty, l'un des pères fondateurs de la formule "2.0 ", fut lui aussi l'un des pionniers du web, et notamment le fondateur de *GNN*, le premier portail financé par la publicité.

1.1.2. Naissance du concept :

C'est au cours de l'été 2004 que le terme de "web 2.0" apparaît pour la première fois : lors d'une séance de brainstorming, Dale Dougherty, co-fondateur de la société d'édition *O'Reilly*, et Craig Cline de *MediaLive*, réfléchissent à des thèmes émergents en vue d'une conférence à venir. Le journaliste John Battelle, co-fondateur du magazine *Wired*, se joint un peu plus tard au duo et la première conférence Web 2.0 a lieu en octobre 2004. L'expression 2.0 est utilisée à ce propos pour marquer l'émergence d'une nouvelle étape dans la courte histoire du web, étape caractérisée par un changement des règles et une modification des

modèles de revenus. Le volet social sera développé plus tardivement par Tim O'Reilly lors de la deuxième conférence Web 2.0 en octobre 2005.

Les nouveaux produits et services associés au web 2.0 vont ainsi repositionner l'internaute au coeur du système. Des techniques de publications simples et souples comme les blogs, les plates-formes collaboratives de type Wiki offrent des possibilités accrues de production, diffusion et consommation des contenus. L'internaute devient ainsi co-créateur de nouvelles applications en ligne, initiant souvent de façon collective ou communautaire de nouvelles formes de relations numériques. Selon l'expression d'Hubert Guillaud⁶, le web devient non plus une aventure technologique, mais bien une aventure humaine.

1.1.3. Les principes de web 2.0 :

Le texte de Tim O'Reilly, paru en 2005 ⁷, sous le titre "*Qu'est ce que le web 2.0 ?*" dégage sept principes clés du Web 2.0, en voici un petit résumé :

- **Le web en tant que plate-forme** : Actuellement, le système informatique est basé sur des logiciels installés sur des machines indépendantes. Le développement des technologies réseau et leur faible coût pour l'utilisateur (la France est leader mondial du haut débit à bas prix) permettent de remettre en cause cette vision du système informatique. Le web peut maintenant être une plate-forme, c'est à dire qu'un site web peut être un outil aussi efficace pour certaines applications qu'un logiciel. Un exemple concret ? Google offre un tableur et un traitement de texte en ligne. Il s'agit d'une application web pure pour laquelle le web est une plate-forme au sens strict du terme.

Les entreprises ou les utilisateurs qui utilisent un logiciel sous licence traditionnels doivent faire face à l'utilisation du logiciel, le mettre à niveau régulièrement et restez au fait des versions plus récentes, et l'élargissement de faire face à des niveaux appropriés de la demande. La plupart des succès du Web 2.0 ne vendons plus de logiciels sous licence, mais plutôt offrir leur logiciel comme un service. Le client final utilise simplement le service via un navigateur. Tous les maux de tête de la course, la maintenance et mise à l'échelle du logiciel et du matériel sont pris en charge par le prestataire de services de façon transparente pour l'utilisateur final. Le logiciel est mis à jour assez fréquemment par le prestataire de services et est disponible 24x7.

- **Exploitation de l'intelligence collective** : La clé au succès du Web 2.0 applications est comment effectivement ils peuvent armer l'information fournie par des utilisateurs. Plus personnalisé votre service, plus vous pouvez apparier un utilisateur au contenu de son choix mieux.

- **dur-à-reproduire les données comme avantage compétitif** : Difficiles à reproduire, unique, grands ensembles de données fournissent un avantage concurrentiel à un de l'entreprise. Web 2.0 est de données et logiciels combinés. On ne peut pas répliquer Craigslist, eBay, Amazon, Flickr, Google ou tout simplement en reproduisant le logiciel. Les données sous-jacente que le logiciel génère de l'activité des utilisateurs est extrêmement précieux. Cette base de données augmente chaque jour, à améliorer le produit tous les jours.

- **La version bêta perpétuelle** : Le Web 2.0 compagnies libèrent leurs produits tôt pour faire participer leurs utilisateurs et pour gagner le feedback important. Ils réitèrent souvent en ayant les cycles courts de version. Ils font participer les utilisateurs tôt dans le processus. Ils équipent l'application pour saisir la métrique importante sur la façon dont un nouveau dispositif est utilisé, combien de fois elle est utilisée, et par qui. Si vous n'êtes pas sûr comment un dispositif particulier devrait regarder et avoir des conceptions de concurrence, exposer un prototype de chacun à différents ensembles d'utilisateurs et mesurer le succès de chacun. Faire participer les abonnées et les laisser décident lesquels elles aiment. En ayant les cycles de développement courts, il est possible de solliciter le feedback des utilisateurs, les changements incorporés tôt du cycle de vie de produit, et la construction ce que les utilisateurs veulent vraiment.

- **Des modèles de programmation plus simples** : Des versions de développement plus simples mènent à une adoption plus large et réutilisent. Concevoir le votre demande de « hackability » et de « remixability » après des standards ouverts, en utilisant des modèles de programmation simples et une structure d'autorisation qui met en tant que peu de restrictions selon les besoins.

- **Le logiciel au-dessus des machines** : Le logiciel n'est plus lié à une plate-forme. toutes les machines consomment le contenu produit par ces logiciels (gsm, agendas

électroniques, GPS, etc...). Le logiciel devient alors une application synthétique utilisant des web services hébergés sur plusieurs plates-formes. Les outils peuvent alors être utilisés sur des terminaux autres que des ordinateurs, les téléphones portables par exemple.

- **Expérience utilisateur riche** : Le succès d'AJAX a alimenté l'utilisation croissante des interfaces utilisateur riches dans le Web 2.0 applications. Le flash d'Adobe/câble et le Microsoft Silverlight sont d'autres solutions de rechange pour créer UIs riche.

1.1.4 Les évolutions par rapport au web 1.0 : le rôle de l'internaute :

En quoi le Web 2.0 se distingue-t-il de la vague précédente ? On a vu que l'une des caractéristiques majeures du phénomène est l'appropriation des contenus et des modes de diffusion par les internautes eux-mêmes, via des applications souples et légères.

Paul Graham évoque le principe de "démocratie" du Web 2.0, c'est-à-dire la superposition au réseau technique d'un réseau humain qui participe à la création et à l'enrichissement de contenus et services. L'internaute deviendrait ainsi "consomm'acteur" : à lui de produire, publier, diffuser, indexer, partager des contenus de tous types.

	Web 1.0	Web 2.0
Leaders du web	Entreprises, marchands	Internautes
Profil de l'internaute	Passif	Actif
Interactivité perçue	Sélection et lecture d'information	Sélection, lecture et publication de données
Unité de recherche	Mot-clé	Tag

Tableau 1 :Du Web 1 vers le Web 2. Source ZDNet

Pour autant, les internautes vont-ils devenir tous acteurs de ce processus ? Selon un article de Charles Arthur dans *The Guardian* 9, pour un groupe de 100 personnes, si l'une crée du contenu, 10 seulement vont interagir en apportant commentaires et améliorations, les autres se contentant d'une consommation passive. Ce constat, obtenu via des retours d'expérience sur des sites collaboratifs, illustre la "règle des 1/10/89 %", selon laquelle plus de 80% des contenus produits proviennent d'un nombre très restreint d'utilisateurs.

Tendance confirmée par les chiffres présentés lors de la *Web 2.0 Expo* en avril 2007 et repris dans *Internetactu 10* : selon la société de mesure d'audience *Hitwise11*, "0,16 % des internautes qui passent sur *YouTube* y placent des films - alors qu'ils étaient encore 0,5 % à le faire en juillet 2006 ; seulement 0,2% des visiteurs de *Flickr* y publient des photos. La seule exception à cette faible participation provient *Wikipedia*, où 4,6 % des visiteurs contribuent à l'enrichissement de semblent en effet voir écrit des textes longs (en un nombre réduit d'interventions), tandis que les éditeurs" les plus actifs, qui forment le noyau des wikipédiens nettoient, corrigent, lecture fondée sur des actions à faible niveau d'exigence (lire, enregistrer un signet, tagger, commenter), de l'intelligence collaborative, caractérisée par des formes de participations plus actives et intenses : écrire, collaborer, modérer... l'encyclopédie". Jimmy Wales, fondateur de *Wikipedia*, estime qu'un nombre restreint de contributeurs et éditeurs font l'essentiel du travail, alors qu'un nombre considérable de lecteurs apportent peu de modifications. Le "wikipédien" Aaron Swartz nuance ce point de vue : "un grand nombre de contributeurs occasionnels semblent en effet voir écrit des textes longs (en un nombre réduit d'interventions), tandis que les éditeurs" les plus actifs, qui forment le noyau des wikipédiens nettoient, corrigent, mais sans apporter énormément de texte".

Le Web 2.0, et tout particulièrement la blogosphère semble suivre la loi de la participation, édictée par Ross Mayfield, président de la société Social Text, sous le nom de "Power law of participation"¹². Cette "loi", basée sur l'observation des pratiques et des seuils de participation aux dispositifs communautaires, constate que la participation s'amenuise à mesure que les formes de participation se complexifient. Le schéma ci-dessous, qui rappelle celui de la Longue Traîne, distingue l'intelligence collaborative, caractérisée par des formes de participations plus actives et intenses : écrire, collaborer, modérer...

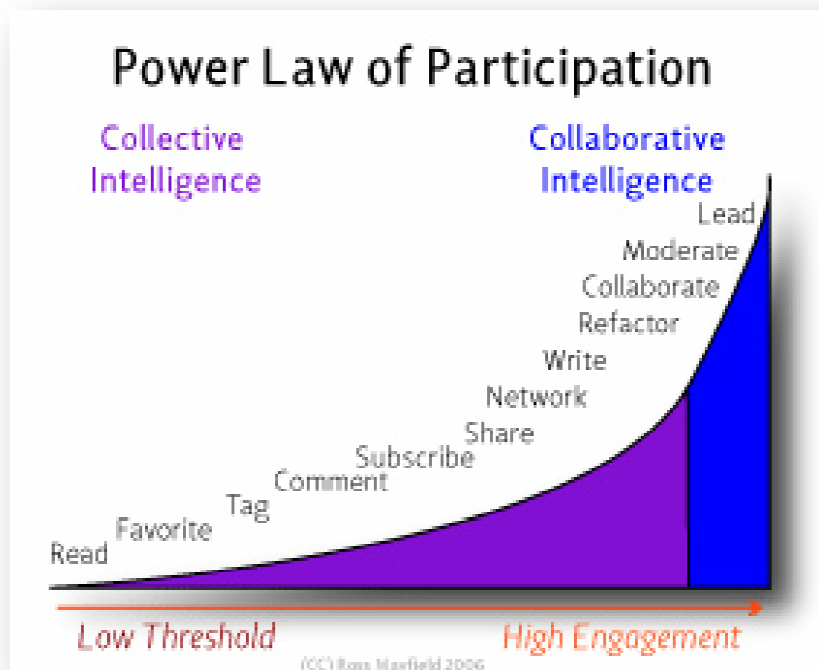


Figure 1.1 : Power Law of Participation, par Ross Mayfield. CC

Le cabinet *Forrester* détaille encde la participation figurant dans son étude *Social Technographics 13* : la grille compte6 échelons correspondant aux niveaux de particip. les inactifs (52 %) . les spectateurs (33 %) : consommateurs de contenus texte ou multimédia (blogs, vidéos mis en ligne par les utilisateurs, podcasts)

- **Les inactifs (52%)**
- **Les spectateurs (33%)** : consommateurs de contenus texte ou multimédia (blogs, vidéos mis en ligne par les utilisateurs, prodcasts)
- **Les minuisiers (19%)** : utilisateurs de réseaux sociaux
- **Les collecteurs (15%)** : utilisent de flux RSS et taggent les pages web
- **Les critiques (19%)** : postent de commentaires sur les blogs, des avis et critiques
- **Les creations (13%)** : publient des contenus en ligne (blog, vidéo..)

1.1.5. Web 2.0 concepts pratiques et technologiques :

La principale caractéristique du web 2.0 est qu'il utilise est basé sur l'utilisation de technologies déjà existantes et sûres. De ce fait il ne constitue pas en soi une révolution technologique mais une succession d'évolutions. Il est à noter que la principale contrainte des

développeurs étant le parc de navigateurs installés sur les machines des internautes, leur souci a toujours été de développer des applications accessibles aux plus grands nombre.

✓ **Structuration de l'information, des données à valeur ajoutée :**

Un ensemble de technologies (telles que celles du "web sémantique" qui enrichit le web de contenu assimilable par la machine et celles associées à l'agrégation-syndication-flux RSS) permettent de relier les informations, d'agréger ou de partager des services et des contenus, de refondre les interfaces, de participer à des espaces collaboratifs, de personnaliser sa pratique de l'Internet et de prendre part plus facilement à son développement.

✓ **Une interactivité toujours plus grande :**

Renouveau des interfaces utilisateurs des sites web. Interfaces riches, Rich media (intégration de la vidéo et de l'audio). Les technologies AJAX et FLASH enrichissent les interfaces utilisateurs de fonctionnalités nouvelles. Ajax aide la fluidité de navigation en évitant des rechargements de pages lorsque l'on modifie certains éléments. Il rend également possible les "glisser-déplacer" d'objets tout comme dans les environnements de type Windows. Le web devient encore plus interactif.

✓ **Interopérabilité :**

Le web 2.0 combine des architectures plus flexibles, des protocoles de communication plus ouverts (Web Services) avec notamment REST et SOAP, une interopérabilité plus poussée, une amélioration des interfaces utilisateurs d'un point de vue ergonomique et sémantique qui donnent des résultats d'une grande souplesse d'utilisation.

✓ **Modularité :**

Les applications et services fournis par divers sites sont autant de briques d'application qui peuvent, grâce aux Web Services les rendant interopérables, être recombinaées à loisir à d'autres fonctionnalités pour composer, créer de nouvelles application, pour proposer de nouveaux services. Ainsi, les (mashups) ajoutent des couches plus pratiques et fonctionnelles au web. L'avantage pour les utilisateurs est immédiat : disposer d'un guichet unique pour consulter différents services.

✓ **Montée en puissance des applications bureautique en ligne :**

Microsoft a bien senti le renversement de situation en cours avec son ambitieux programme Office Live, ainsi que Google, avec sa plateforme en ligne "Google Docs": l'avenir n'est plus aux applications autonomes installées sur le poste des utilisateurs mais plutôt aux

applications en ligne qui seront facturées à la demande. Ainsi, de nouveaux services en ligne ont fait leur apparition (Suite bureautique, traitement de texte, tableur, création de diagrammes, client email, gestion de calendrier, gestion de projet).

A terme, les services de Web 2.0 pourraient remplacer les applications bureautiques. Les logiciels utiles se libèrent d'une plate-forme spécifique. Ils deviennent accessibles partout et sur tout support. Vers plus de mobilité.

✓ **Une mutation perpétuelle des services :**

La version beta d'un service n'est plus une étape vers une version figée : elle devient la situation "stable" d'un service en changement dynamique. Elle est mise en ligne alors même qu'elle est en cours d'évolution. Cela permet de générer du buzz autour de ses nouveaux services, d'avoir des bêta testeurs, de faire évoluer le service et ses infrastructures petit à petit (grâce notamment à l'analyse du comportement des internautes face aux nouvelles fonctionnalités proposées au fur et à mesure).

✓ **Une démarche open source à travers le développement participatif :**

Une grande partie de l'infrastructure du web elle-même repose sur les méthodes de production individuelle de l'open-source qui sont, elles aussi, une émanation de l'intelligence collective. Les utilisateurs deviennent des co-développeurs.

1.1.6. Glossaire du Web 2.0 :

Pour mieux comprendre les concepts utilisés dans l'univers du Web 2.0, nous vous proposons ce glossaire des termes essentiels :

- **Tags** : mot clef sous forme d' "étiquette" que les utilisateurs peuvent apposer sur un document numérique (texte, photo, vidéo), de façon à en décrire le contenu. Le *tagging* introduit une part de subjectivité par rapport à l'indexation automatique effectuée par les moteurs. Les tags peuvent également être extraits automatiquement d'un texte en fonction de leur fréquence d'occurrence.

- **Nuage de tags** : présente les tags correspondant à un site (ou à plusieurs) en les regroupant sur un même espace et leur attribuant une taille variable suivant la fréquence d'utilisation ou d'occurrence. Cela donne une vision d'ensemble des principales thématiques. Le site *Tagcrowd.com*¹⁴ permet de générer un nuage de tags à partir d'un texte.

- **Folksonomie** : (traduit également parfois par potonomie, peuplonomie...) Ce mot valise, contraction de folk (peuple) et taxonomie, désigne une forme de "classification collaborative décentralisée spontanée" basée non pas sur un vocabulaire contrôlé et standardisé mais sur des termes choisis par les utilisateurs eux-mêmes, intitulés tags. L'objectif étant de faciliter l'indexation des contenus, ainsi que la recherche d'information.

- **Blogs** : journal personnel disponible sur le web. Peut être tenu par un particulier, un chercheur, un journaliste, un salarié d'entreprise ou plus rarement par un groupe de personne (entreprise, collectivité...). Le blog a pour objet de diffuser des billets, généralement liés à l'actualité, et présentés par ordre chronologique, alors qu'un wiki est destiné à la mutualisation des connaissances. Un blog peut présenter des contenus multimédias : photo (photoblog), son ou vidéo (vblog, vlog)...

- **Blogosphère** : la "biosphère des blogs ". Ce mot valise désigne la communauté des blogs, ou encore l'ensemble des auteurs de blogs. La blogosphère est largement interconnectée, les trackbacks situés sur la plupart des blogs permettant de lier deux articles traitant de sujets connexes. L'élite des blogueurs particulièrement influents peut être qualifiée ironiquement de "blogeoisie".

- **Blogroll** ou **blogoliste** : liste de liens vers d'autres blogs, présentés par l'auteur d'un blog. On peut syndiquer sur une même page des billets venant de ces blogs via les formats RSS.

- **RSS**: l'acronyme signifie selon les versions *Really Simple Syndication* (RSS 2.0), *RDF Site Summary* (RSS 0.9, 1.0 et 1.1) ou encore *Rich Site Summary* (RSS 0.91). Ce format de syndication permet d'extraire automatiquement d'un site web ou d'un blog du contenu régulièrement mis à jour. Un fichier RSS est un simple fichier texte au format XML comportant la description synthétique du contenu. Ce format a été créé en 1999 par Netscape. On peut s'abonner gratuitement à plusieurs flux RSS à partir d'un lecteur on-line ou installé sur le PC.

- **Atom** : Format normalisé de syndication de contenu Web, basé sur le XML, qui permet, la syndication de contenu entre différentes ressources Web. Concurrent de RSS, Le format Atom a été créé dans le but de simplifier les différents formats RSS existants. La

différence entre les deux formats réside surtout dans l'emploi de balises XML différentes. Les deux formats sont généralement lus par les agrégateurs.

- **Wiki** : Site web dynamique dont tout visiteur peut modifier les pages à loisir. Le terme vient la langue hawaïenne (wikiwiki = rapide). Un wiki a pour objectif de mutualiser des connaissances et est organisé autour d'un plan de classement. Les contributeurs d'un wiki sont généralement anonymes.

- **Crowdsourcing** : Littéralement "approvisionnement par la foule". Le *crowdsourcing* consiste à utiliser la créativité, l'intelligence et le savoir-faire d'un grand nombre d'internautes pour créer des contenus, participer à l'alimentation d'un site. Cette participation des utilisateurs à la création des contenus représente une forme d'externalisation (outsourcing) et permet de réduire les coûts en se basant sur un système de bénévolat ou de micropaiement.

- **Réseaux sociaux** : de nombreux sites de "social networking" ou réseautage social sont apparus avec la vague du Web 2.0. Ils permettent de mettre en relation des personnes partageant les mêmes centres d'intérêt personnels ou professionnels. Ces réseaux peuvent devenir de véritables communautés en ligne, les plus fameux étant les sites *Friendster* ou *Myspace*, ou bien dans la sphère professionnelle *Linkedin*, *Viadeo*...La plupart des réseaux sont ouverts, certains fonctionnent sur le mode de la cooptation.

- **Podcasts** : issu de la combinaison des termes *iPod* (d'Apple) et *broadcasting*, ce terme désigne le fait de rendre disponible en ligne un fichier audio au format numérique. Ce fichier peut-être téléchargé directement sur un ordinateur ou un périphérique, via un flux de diffusion. Equivalent français : baladodiffusion. L'utilisateur peut ainsi se composer une liste de lecture à la carte. Pour les fichiers vidéo, on emploie le terme de vidéocast.

- **AJAX** (Asynchronous JavaScript And XML) : méthode informatique de développement d'applications Web permettant d'économiser de la bande passante, en ne rechargeant pas une page entière alors que seuls certains éléments ont besoin de l'être, mais en ne rafraîchissant que ces éléments de la page. Cela permet de produire des contenus dynamiques et réactifs en ne nécessitant que peu d'appels au serveur hébergeant la page. AJAX est une combinaison de plusieurs technologies : HTML, CSS, Javascript, XML...De nombreux sites web 2.0 sont développés en AJAX.

- **API** : Interface de programmation permettant à un utilisateur de développer des applications conviviales, De nombreux sites classiques ou "Web 2.0" proposent des API pour intégrer automatiquement leur contenu à d'autres sites.

- **Mashup** ou **Mash up** : application web composite mixant plusieurs sources ou plusieurs contenus pour fournir un nouveau produit ou service. Le terme se traduit en français par remixage, ou mosaïque. Exemple de mashup : la visualisation des résultats des élections présidentielles françaises via Google Earth qui mixe les cartes Google et les chiffres du Ministère de l'Intérieur.

- **Long tail (longue traîne)** : Cette expression est utilisée en statistiques, et peut s'appliquer au marketing. Des produits rares ou qui font l'objet d'une faible demande, peuvent collectivement représenter une part de marché égale ou supérieure à celle des best-sellers, si les canaux de distribution peuvent proposer assez de choix. Le caractère virtuel du web permet de générer une longue série de marchés de niche hétérogènes (exemple : *e-bay*), par opposition aux grands marchés homogènes et grand public du monde réel.

1.2 L'intelligence collective et le web 2.0 :

C'est vraiment le cœur du Web 2.0, c'est l'intelligence collective, que j'ai défini comme exploitant l'effet de réseau pour créer des applications aller mieux plus les gens les utilisent." Ce sont les paroles de *Tim O'reilly* à l'Expo Web 2008 .

L'intelligence collective est la capacité des communautés d'évoluer vers une plus grande complexité des commandes et l'harmonie, l'innovation grâce à des mécanismes tels que les variations de rétroaction-sélection, la différenciation-intégration –transformation, et la concurrence-coopération-coopétition.

Le principe central du succès des géants nés du web 1.0 qui ont survécu pour maintenant mener l'ère du web 2.0, semble être qu'ils ont su exploiter de manière efficace la force de l'intelligence collective que recèle potentiellement le web :

- Les liens hypertextes sont le fondement du web. Au fur et à mesure que les utilisateurs ajoutent des contenus, celui-ci est intégré à la structure du web par d'autres utilisateurs qui les découvrent et placent des liens vers ceux-ci. Telles des synapses formant un cerveau où les associations se renforcent à force de répétitions et d'intensité, les connections au sein du web se multiplient organiquement à la mesure que leur dicte l'activité de l'ensemble des utilisateurs.

- Yahoo !, la première grande success story, est né d'un catalogue de liens, une agrégation du meilleur de milliers puis de millions d'internautes. Même si Yahoo! a depuis quelque peu changé de modèle d'affaire, son rôle en tant que portail vers l'oeuvre collective des utilisateurs du web reste sa principale valeur.

- La percée de Google dans la recherche, qui en a fait rapidement et indiscutablement le maître du marché, était basée sur le PageRank(*ndt : littéralement classement de page*), une méthode basée sur la structure hypertexte du web plutôt que sur l'analyse des contenus.

- Le produit d'eBay est l'activité collective de l'ensemble de ses utilisateurs ; comme le web lui-même, eBay grossit organiquement en réponse à l'activité de ses clients, le rôle de la société consistant simplement à fournir le contexte dans lequel cette activité peut se dérouler. De plus, les avantages comparatifs d'eBay viennent entièrement de la masse critique d'acheteurs et de vendeurs disponibles, tout nouveau concurrent étant inévitablement pénalisé de ce point de vue.

- Amazon vend les mêmes produits que des concurrents tels que Barnesandnoble.com, ils reçoivent les mêmes descriptions de produits, images de couverture et contenus de la part des éditeurs. Mais Amazon a fait une science de l'implication du consommateur. Ils sont un cran au dessus des autres en ce qui concerne les commentaires des acheteurs, les invitations à participer (de plusieurs manières et sur quasiment toutes les pages) - et plus important encore, ils se servent de l'activité des utilisateurs pour affiner les résultats de leurs recherches. Alors que les recherches de Barnesandnoble.com poussent les utilisateurs vers ses propres produits ou vers des résultats sponsorisés, Amazon renvoie toujours vers « les plus populaires », une notion calculée en temps réel non pas à partir des ventes mais à partir de ce que les gens d'Amazon appellent le « flow » (*ndt : le bruit, le mouvement, le flux*) autour des produits. Il n'y a donc rien d'étonnant à voir la croissance des ventes d'Amazon dépasser celles de ses concurrents.

1.3. les nouveautés du web 2.0 :

Le web 2.0 est le fruit de plusieurs changements technologiques, sociales, culturels, économiques du réseau Internet .Il n'est pas une nouvelle version du web comme c'est le cas pour le logiciel où l'ancienne version est dépassée mais plutôt une évolution de ce dernier. Il possède sans doute d'énormes avantages mais ne manque pas de limites et inconvénients. Certains avantages de ce type de web peuvent devenir des inconvénients. C'est la raison pour laquelle on propose dans ce qui suit de présenter les nouveautés du web 2.0 par rapport au web 1.x et de les discuter.

1- Premier changement : Participation des utilisateurs

C'est que le web 1.x était un web où un producteur produit pour plusieurs consommateurs. Maintenant les utilisateurs sont devenus producteurs et consommateurs. Donc au lieu de consommer seulement on peut participer. Autrement dit si le web 1.x c'est pour tous, maintenant c'est tous pour tous. En effet, dans le web 1.x le concepteur du site était la personne qui se charge de la production du site et de sa mise à jour maintenant tout le monde peut créer son site sans aucune restriction. Dans cette nouvelle optique, chaque page est une collaboration des auteurs. Donc tout le monde peut avoir sa place dans ce web 2.0 (professionnels, entreprises, débutants,..). C'est une démocratisation de l'information et une liberté des avis qui permet le partage , collaboration, échange, suivis, retracé...

Ce mode favorise le travail collectif, collaboratif et participatif qui cumule les connaissances humaines et qui permet une forte capitalisation et Co-construction simplifiée toute en s'adaptant à une large audience. Il favorise la mobilisation de l'intelligence collective et le travail en équipe sans contrainte géographique ni politique ni imposer aucune restriction.

Reste à voir qui produit et quand et quel est l'âge et le niveau intellectuel de celui qui produit. Donc certainement la quantité d'information va augmenter mais au détriment de la qualité.

2- Deuxième changement : Large gamme des outils

Contrairement au web 1.x, le web 2 est simple et facile à utiliser, Il est devenu une plateforme de services et non plus un regroupement de sites. Pour cela il offre des outils de communication et de partage de l'information partout et par conséquent des services sur le web à travers diverses applications (Les blogs, les wikis, les réseaux sociaux, les flux Rss,...). Ces outils sont d'abord centrés sur les utilisateurs qui sont également leurs contributeurs et leurs bénéficiaires. Ils sont simple et faciles à utiliser et ne demande pas de formation ; Ils permettent une large diffusion des informations et s'interfacent très facilement et rapidement avec d'autres outils ce qui se traduit par un gain de temps, d'effort et de cout. Par exemple avec les flux RSS on peut faire de la veille. L'utilisateur peut souscrire à une page web et d'être notifiés de tout changement. Avec les Tags, on peut faire de la bonne recherche mais les redondances rendent cette possibilité trop couteuse. Le Web 2.0 peut être un outil d'apprentissage et par conséquent un outil de formation continue mais qui peut toujours subir des améliorations.

3 - Troisième changement : Portabilité et programmation souple et légère

Poussé par les avances technologiques au niveau des supports de stockage, du haut débit et aussi l'évolution des protocoles web, on ne parle plus de texte et photo maintenant mais du son et vidéo (Blog : Audioblogs, podcast, vidéoblogs). La taille des sauvegardes a migrée du MO vers le GO voir même les TO. Le web 2.0 est devenu très répondu grâce à la baisse des prix et encore plus rapide grâce au haut débit. Il en découle un nombre important d'internautes repartis sur toute la terre et plus spécialement l'intégration des pays du tiers monde en réduisant la fracture numérique. Par ailleurs, l'utilisation de l'AJAX qui est un langage léger permet d'étendre les possibilités de codage dans les navigateurs ce qui se traduit par des pages conviviales et très interactives. Le Web 2.0 permet également de diversifier les terminaux grâce aux versions d'Ajax sur mobile à travers leurs navigateurs et a l'apparition des réseaux mobiles (réseaux wifi, wimax et 3get 4g). Ce qui permet d'avoir le privilège d'utiliser des machines moins puissantes en occurrences les mobiles et les iphones. Comme exemple de l'apport du web 2.0, on peut citer l'encyclopédie Wikipédia qui a demandé des années seulement pour devenir le wiki le plus répondu et la référence pour les chercheurs et

les utilisateurs sans payé les frais des contenus contrairement a beaucoup de sociétés qui ont dépensé beaucoup de temps et d'argent.

1.4. Les limites du web 2.0 :

Comme on a déjà dit, le web 2.0 possède des avantages mais aussi des limites. Sur la base de nouveautés du web 2.0, on peut dégager beaucoup de limites.

1- Limite 1 : Mauvaise qualité de l'information

Grâce au Web 2.0, tout le monde peut: créer, publier, partager, se connecter, influencer, collaborer; Oui mais quelle est la qualité de ce qui est publié et sa pertinence ? D'après ma propre expérience en tant qu'enseignant chercheur, dans la même classe, il ya toujours des étudiants brillants ceux moyens et enfin ceux qui sont très limites. Donc la participation ne sera jamais la même. Puis il y a ceux qui veulent participer et ceux qui ne le veulent pas. Pour participer, il faut au moins des connaissances et compétences minimales avant d'aborder un sujet pertinent. Ce qui s'applique exactement sur le web 2.0, les gens n'ont pas la même formation, les mêmes expériences, les mêmes idées. Donc, parmi les limites du web 2.0, c'est qu'aucune vérification réelle des gens n'est faite ce qui donne une possibilité de mensonges ou de mauvaise utilisation d'autant plus qu'on peut utiliser de faux pseudos. Le mode actuel du web 2.0 donne beaucoup d'informations redondantes, inutiles et non intéressantes ce qui nous oblige à le réadapter.

2- Limite 2 : Faible taux de participation

Le web 2.0 instaure une nouvelle démocratisation de l'information qui donne l'accès de lire et de participer à tout le monde. Oui mais est ce que tout le monde publie ou au moins la majorité. selon plusieurs études , peu d'utilisateurs le font, en effet 1% des utilisateurs produisent 2 /3 de la production . En 2004, Par exemple, Wikipedia à un total de 68.682 internautes 80% de toutes les modifications de Wikipédia sont effectués par 10% des utilisateurs et 5% des utilisateurs ont écrit plus de 66% de tous les articles . De plus ce nombre est en décroissance totale .Donc avec ce nouveau mode il y a beaucoup de lecteurs oui mais pas beaucoup de participants d'entre eux .

3- Limites 3: Durée de vie de l'information très courte

Parmi les problèmes que pose les applications du web 2.0 existe le problème de la mise à jour, en effet les blogues et les wikis doivent être alimenté régulièrement pour ne pas perdre leurs utilisateurs. L'information qui contient est souvent inactive. Les blogues se créent et meurent rapidement à cause du taux d'abandon qui est très élevé. L'information n'a plus de valeur et ne s'archive sur les serveurs même si elle est très intéressante, en fait il faut stocker l'information non pas par l'ordre inverse de l'ordre chronologique mais plutôt par l'ordre d'importance.

4- Limite 4 : Manque de sécurité et droit d'auteurs

Les droits d'auteurs posent un grand problème pour le web 2.0. Les internautes bénéficient de l'information existante sur les commentaires postée sur un blog par exemple. Il peut s'opposer à toute reproduction sans son autorisation ou au moins exiger qu'elle soit mentionnée. A mon avis ce problème ne doit même pas être posé pour le résoudre dans les tribunaux, comme exemple on peut citer le problème de youtube et les droits d'auteurs des vidéos qui le composent. Cette société a donné aux tribunaux en USA les coordonnées des collaborateurs et a créé un grand problème qui ne doit même être posé. En plus des droits d'auteurs existe les Spam et les malfaiteurs qui n'arrêtent pas de nous embêter et ne pas nous laisser profiter de cette collaboration et partage du web 2.0.

5- Limite 5 : Manque de sémantique

Le Web sémantique a été proposé en premier lieu par l'inventeur du World Wide Web, Tim Berners. Le réseau web 2.0 n'est pas un réseau sémantique. En effet, l'information n'a pas de sens ou d'interprétation et se reflète par la non possibilité d'effectuer des recherches très pertinentes qui intéresseront grandement les spécialistes. Les données des contenus peuvent être enrichies afin de leur donner un sens afin d'enrichir et de personnaliser les résultats de recherche associés à ces contenus. L'ajout de sens afin d'augmenter la pertinence des contenus et il faut penser à la constitution d'une base de données à l'échelle du web.

1.5. Pourquoi le Web 2.0 marche-t-il ?

Le Web 2.0 fonctionne pour les raisons suivantes. Les métadonnées sont encore des données, quand bien même il s'agit de données sur des données ; c'est-à-dire que ce sont des différences identifiables qui fournissent et contraignent une interprétation sémantique mais qui en sont encore dépourvues par elles-mêmes (Floridi 2008A). Ces données ne doivent pas être confondues avec des informations sémantiques (qui nécessitent au moins la signification, Floridi 2005), sans parler de la connaissance (qui nécessite certainement la vérité et peut exiger en outre au moins une forme de justification ainsi que la compréhension).

Cependant, les agents artificiels – y compris les logiciels quotidiens et, dans la mesure où les connaissances scientifiques actuelles sont en jeu, tout logiciel concevable qui peut être développé à l'avenir - sont des *moteurs syntaxiques* qui ne peuvent pas traiter des données pourvues de signification, c'est-à-dire des contenus informationnels, mais seulement des données à un niveau faible ou supérieur. Ainsi, le Web sémantique est largement un simple battage : nous avons vu, par exemple, que XML est un langage de description de données où aucune information n'est ou ne peut être impliquée.

Les humains au contraire sont les seuls moteurs sémantiques disponibles, les fantômes dans la machine. Ainsi le Web 2.0 est le Web créé par des moteurs sémantiques pour des moteurs sémantiques, en s'appuyant sur la contribution de légions d'utilisateurs. À titre d'exemple, considérons les folksonomies.

Une folksonomie (mot construit à partir de folk et de taxonomie) est le résultat cumulé de la pratique sociale de production d'informations (de métadonnées, pour être précis) portant sur d'autres informations (par exemple, une photographie) par le biais d'une classification collaborative connue sous le nom d'étiquetage ou de tagging social (par exemple, la photo reçoit les tags "New York", "Hiver", "Statue de la Liberté"). Elle fonctionne de bas en haut, puisque est laissé au seul utilisateur individuel ou au producteur de la cible ainsi "tagguée" le choix de ce qui à classer, comment le classer et quels mots-clés utiliser dans la classification. Les folksonomies sont devenues populaires depuis 2004 en tant que moyen efficace de personnaliser l'information et de faciliter sa concrétisation grâce à des outils de gestion de l'information. Il est maintenant presque trivial de remarquer que les folksonomies peuvent être

extrêmement ambiguës, mais ce n'est pas un problème pour les moteurs sémantiques que nous sommes, capables de processus rapides de résolution des ambiguïtés.

Il s'avère que le Web 2.0 constitue un objectif accessible et de plus en plus effectivement réalisé, représenté non plus par la création d'un autre espace extérieur, à l'instar du Web 1.0, mais par un écosystème bienveillant et habité par les humains considérés comme des inforgs (c'est-à-dire des organismes informationnels connectés).

L'explication précédente clarifie le fait que le Web 2.0 soit une partie d'un espace d'information, *l'infosphère*, où la vue platonicienne de la mémoire conçue comme enregistrement et préservation éternelle est remplacée par la vue de la mémoire conçue comme accumulation et perfectionnement.

Il s'agit d'un environnement caractérisé par le fait qu'il est *favorisé par le temps* : le temps ajoute de la valeur et les applications Web 2.0 deviennent meilleures à l'usage, elles s'améliorent avec l'âge, notamment parce que le nombre de personnes impliquées augmente constamment. Ceci à son tour est fonction de la masse critique des "prod-umers" [producers/consumers], c'est-à-dire des producteurs et consommateurs d'information sémantique que j'ai défini plus haut comme des inforgs. Les articles de Wikipedia par exemple deviennent meilleurs plus longtemps ils sont en ligne et plus ils sont utilisés 4, ceci également parce qu'une nouvelle génération de participants toujours plus nombreux intensifie l'effet de révision. Cela explique pourquoi le Web 2.0 est considéré comme faisant partie d'une évolution plus récente connue sous le nom de *Cloud computing*. Il s'agit là d'une autre métaphore et d'un mot à la mode pour désigner l'Internet de façon également assez floue et imprécise. Toutefois, comme dans le cas du Web 2.0, l'expression *Cloud computing* renvoie à un véritable nouveau paradigme quand elle est utilisée pour désigner une version plus évoluée de l'utilisation de ressources de calcul : outils logiciels, espace mémoire, puissance de calcul et autres services ou capacités informatiques sont tous fournis en permanence comme des services basés sur Internet (dans le "nuage") d'une manière entièrement transparente en ce qui concerne l'infrastructure et également transparente pour l'utilisateur. Il s'agit là de l'ultime défi à la localisation spatiale et, de ce fait, à la fragmentation des traitements de l'information. Le *Cloud computing* constitue un espace convivial en ce sens qu'il ne dépend pas du lieu où vous êtes mais seulement de ce dont vous avez besoin comme ressources informatiques.

Le Web 1.0 et le Web sémantique, au contraire, sont défavorisés par le temps et ne comptent pas sur le grand nombre de petites contributions qui peuvent être offertes par des

millions d'inforgs. Par exemple, plus un article de la Britannica est disponible depuis longtemps, pire il est, jusqu'à devenir totalement dépassé à longue échéance ; il en est de même pour les vieux sites web qui fonctionnent de façon centralisée, ou de toute ontologie. Ainsi, un test simple pour savoir si quelque chose appartient au Web 2.0 consiste à se demander s'il s'améliore avec le temps, avec son usage, et par conséquent avec le nombre de personnes connectées. Les services qui passent avec succès ce test sont Flickr, YouTube, MySpace, etc.

1.6. L'évolution du Web : la défragmentation de l'Infosphère :

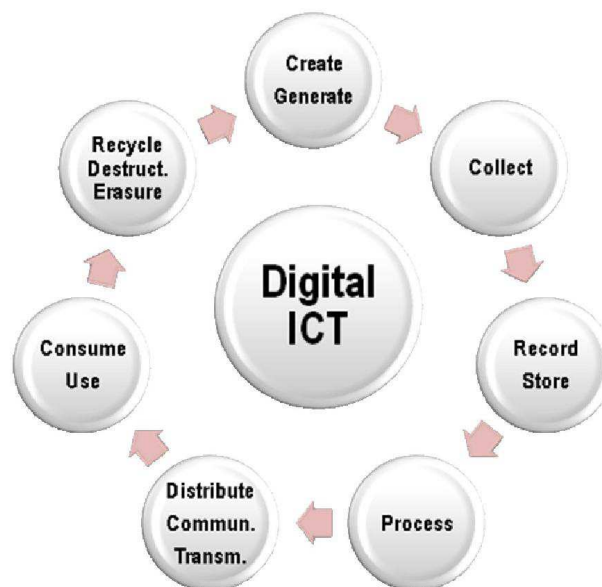


Figure 1.2 : Le cycle de vie de l'information

Pour résumer, la totalité du Web sémantique est une erreur bien définie, tandis que le Web 2.0 est une réussite mal définie. Ils constituent tous les deux des cas intéressants d'un phénomène plus vaste qui peut être caractérisé comme la *construction et la défragmentation de l'infosphère*. Le Web 2.0, c'est-à-dire le Web participatif, efface les barrières entre la production et la consommation de l'information en introduisant moins de friction au sein d'une ou de plusieurs phases du cycle de vie de l'information (depuis sa production jusqu'à son utilisation en passant par son traitement et sa gestion, (voir la Figure 1.2), ou efface les

barrières entre les producteurs et les consommateurs de l'information. Le Web 3.0, le Web sémantique, compris comme il devrait l'être c'est-à-dire comme un Web MetaSyntactique, efface les barrières entre les bases de données. Nous pourrions alors qualifier le Web qui effectue le pont de Web 4.0, qui efface la fracture numérique entre celui qui est et celui qui n'est pas citoyen de la société de l'information au regard de la disponibilité et de l'accessibilité. Fait intéressant, cela se passe plus en termes de téléphones intelligents et autres appareils portables - par exemple en Chine et en Inde - qu'en termes de commercialisation des ordinateurs personnels. Le Web 5.0 ferait référence au *Cloud computing* et à sa capacité à effacer les barrières physiques et à globaliser le local. Enfin, le Web 6.0 serait le Web intégré à la vie (Onlife), qui efface le seuil entre ici (le hors-ligne, l'analogique, basé sur le carbone) et là-bas (le en ligne, le numérique, basé sur le silicium). Dans ce dernier cas, d'autres qualificatifs sont envisageables tels que "l'informatique omniprésente", "l'intelligence ambiante", "l'Internet des objets" ou "les objets augmentés par le Web". Tous ces différents Web se développent en parallèle et, de ce fait, ne figurent qu'en partie selon leur ordre chronologique d'apparition. Leur numérotation n'implique pas d'ordre hiérarchique, mais constitue seulement une question pratique d'étiquetage. Ils devraient plutôt être considérés comme des forces convergentes qui poussent l'évolution du web dans le sens d'une meilleure infosphère.

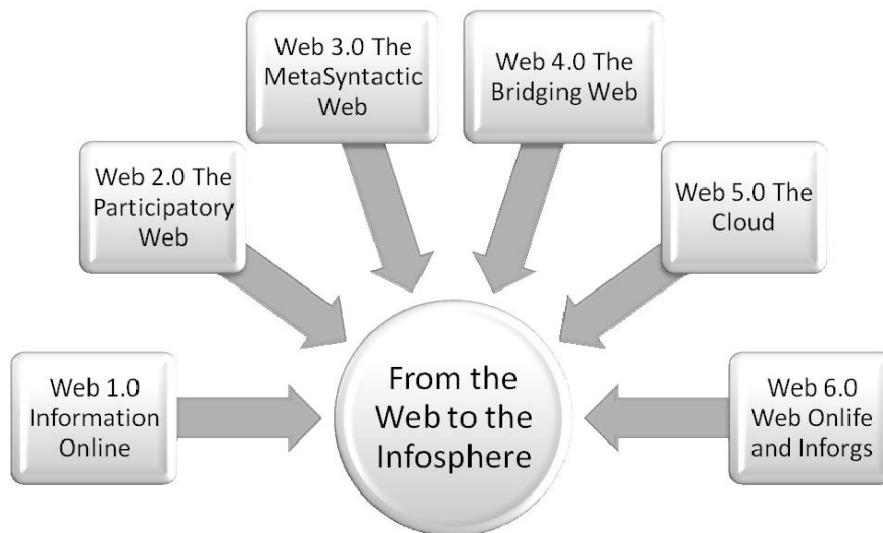


Figure 1.3 : Cartographie de l'évolution du Web

1.7. Les web sémantiques :

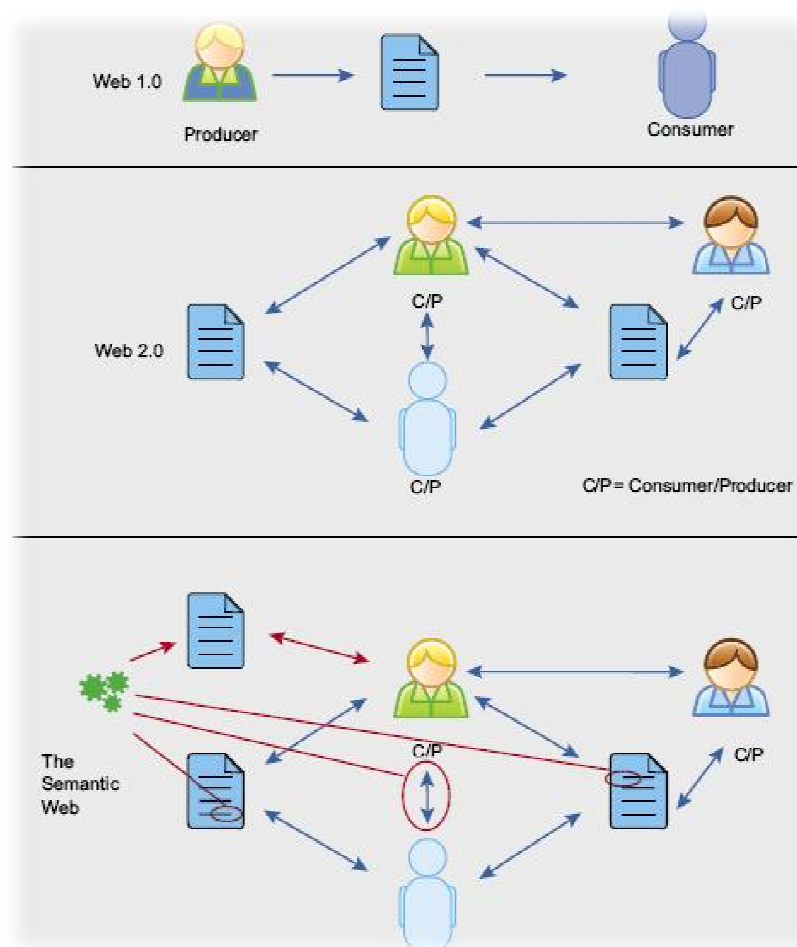


Figure 1.4 : Web 1.0 et 2.0 comparés au Web Sémantique

L'Internet est aujourd'hui une gigantesque base de données très riche mais absolument pas organisée. Le seul soupçon d'organisation visible est sa structure en graphe orienté (généré par les liens hypertextes).

Pour s'y retrouver, il existe les annuaires (de moins en moins utilisés) et les moteurs de recherche.

Ces derniers sont utilisés à longueur de journée pour trouver la moindre information. Ils apportent la plupart du temps la réponse à notre question quand elle est simple. Cependant ils présentent tous certains défauts :

- Ils ne gèrent pas les homonymes. Par exemple, rechercher *chat* sur Google renvoie une liste de sites, certains concernant l'animal et d'autres les salons de discussions.

- Ils ne gèrent pas non plus les synonymes. Si je cherche *achat automobile*, je suis aussi intéressé par les sites qui contiennent le mot *voiture* à la place d' *automobile*, mais aucun ne me les donnera dans les résultats.

- Ils ne savent pas interpréter le langage naturel. Si je demande à Google *quand Bill Gates est-il né ?*, il tentera de trouver des sites contenant le plus possible d'expressions contenues dans la question.

Mais cela ne donne aucune réponse pertinente. Comme nous ne savons toujours pas apprendre une langue naturelle à un ordinateur (cela viendra peut-être un jour), c'est une fois de plus à l'homme de s'adapter.

Le Web sémantique (Figure 1) consiste à ajouter des informations cachées destinées à être utilisées par des applications, des robots de moteurs de recherche, etc.

Ces informations sont généralement présentées sous LE format du Web sémantique, le RDF. Pour faire simple, le RDF, basé sur le XML, permet de structurer des données sous la forme de triplets {Sujet, Prédicat, Objet} :

- Le sujet représente la structure à décrire,
- Le prédicat représente un type de propriété applicable à cette ressource,
- L'objet représente une donnée ou une autre ressource.

Simplement avec des informations sous la forme de triplets, beaucoup de problèmes actuels des moteurs de recherche peuvent être résolus : définir facilement des synonymes, des catégories...

Ces triplets sont écrits directement sur les pages Web. Si des moteurs de recherche les utilisent, nous pouvons imaginer la possibilité de répondre à des questions de manière bien plus précise qu'aujourd'hui.

Seulement, aujourd'hui, les triplets manquent à l'appel. il y a du RDF un peu partout, mais il est utilisé la plupart du temps en tant que micro-format. Peu nombreux sont les sites qui publient du RDF avec du contenu de qualité. Cependant, le Web sémantique représente un intérêt considérable pour les entreprises.

Même si les données RDF sont encore peu présentes, certains se sont penchés sur le sujet soit en essayant de combler cette lacune soit en se dispensant de l'utilisation du RDF.

Conclusion :

Le web 2.0 et tout son apanage technique et technologie apparaît logiquement comme novateur dans sa modification fondamentale de la communication sur Internet. Le serment du Web 2.0 est donc d'ouvrir une ère nouvelle où les utilisateurs sont à la fois les contributeurs et les bénéficiaires. Dès lors la perception de l'intelligence collective prend tout son sens. Un point de départ où une multitude de nouveaux usages et de nouvelles applications sont encore à trouver.

Il n'en demeure pas moins que cette notion conceptuelle qu'est le Web 2.0 reste floue à définir mais aussi à cerner... Pour le moins paradoxal en soit puisque tout le monde en parle mais pourtant personne en saurait en donner une définition exacte sur laquelle tout le monde s'accorderait. Il s'agit surtout d'un ensemble d'enjeux économiques, de nouveaux services définis pour mais aussi autour de l'internaute.

Beaucoup de services estampillés 2.0 ont vu le jour durant ces derniers mois et années, beaucoup sur un même marché d'autres ont su créer la différence et inventer un service apprécié des utilisateurs mais jusqu'à lors inexistant. Cependant, si l'on prend le modèle du web 1.0, première bulle, mais aussi en s'appuyant sur les modèles économiques 2.0 basés en majeure partie sur la publicité ou sur la valeur d'une potentielle entreprise, on peut s'attendre à voir disparaître beaucoup de ces jeunes « start-up » dans un futur relativement proche tant le marché est en pleine évolution devient hautement concurrentiel.

Ainsi en termes d'évolution du Web2.0, on peut imaginer plusieurs choses. Premièrement, l'avènement du concept de Web sémantique va révolutionner la manière de faire ses recherches sur Internet. Deuxièmement, on peut concevoir une évolution de la syndication des données en combinant données et services afin de permettre de développer encore plus le concept de plateforme client/serveur où n'importe quelle application devient accessible via Internet. Enfin, si le terme de Web 3.0 doit exister un jour, cela ne pourra être qu'en tant qu'une évolution du Web2.0 avec des services encore plus poussés, plus intuitifs et sûrement le principe des MashUps exploités au maximum avec des sites concentrant un ensemble de services.

Chapitre II

Le tagging collaboratif

Introduction :

L'organisation efficace de l'information est apparue comme une difficulté majeure dès le début d'Internet. La quantité énorme des informations disponibles impose le recours à des solutions automatisées comme les moteurs de recherche (e.g., Google) ou les annuaires en ligne (e.g., Yahoo). Aujourd'hui, les contenus disponibles sur le Net augmentent notamment sous l'effet de la participation de plus en plus active des internautes qui deviennent créateurs de textes, d'images, de vidéos, etc. qu'ils partagent sur la Toile. Cette tendance se manifeste par la multiplication des *blogs* et autres espaces personnels. Ces nouvelles pratiques ont été accompagnées par l'apparition de nouveaux modes de navigation et d'organisation des contenus, dont le tagging collaboratif en ligne. Des millions d'internautes l'utilisent quotidiennement sur des sites tels que Flickr, Delicious, MyFeedz, ou Technorati. Le tagging collaboratif décrit le processus par lequel de nombreux utilisateurs ajoutent des métadonnées sous forme de mots-clés à partager contenu. Récemment, le tagging collaboratif s'est développée dans la popularité sur le web, sur des sites qui permettent aux utilisateurs de tagger des signets, photos et autres contenus.

2.1. Tagging collaboratif :

Tagging collaboratif est une pratique qui consiste, pour les internautes, à indexer des contenus en ligne à l'aide de mots clés appelés *tags*. Lorsque plusieurs internautes réunis par un site communautaire tagging les mêmes contenus, une sorte de classification émerge. C'est ce que l'on appelle une folksonomie. Elle permet de faciliter les recherches ultérieures de contenus grâce à une représentation visuelle sous forme de nuage de mots. Ces différents concepts méritent de plus amples explications.

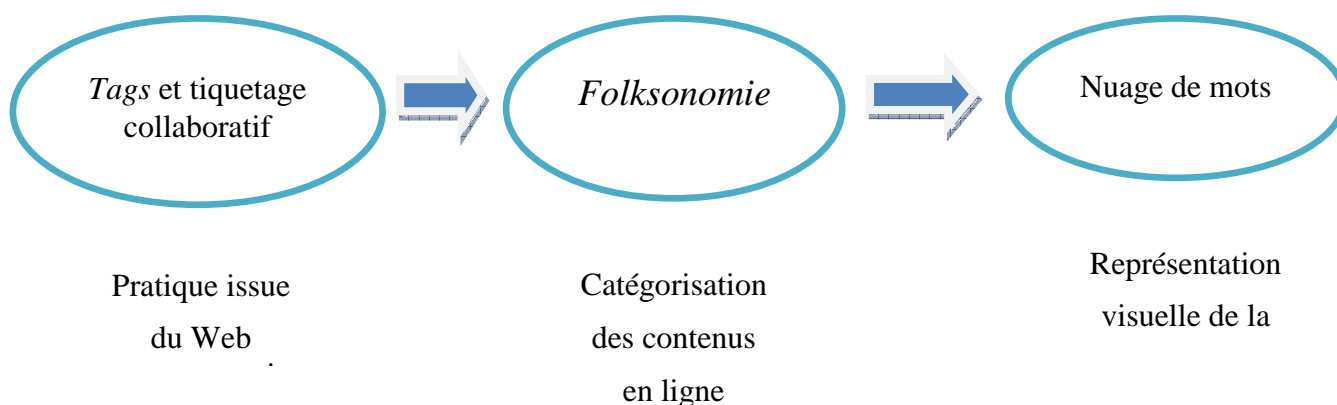


Figure 2.1 : Les principaux concepts

Tagging du contenu par des termes descriptifs, aussi appelé mots-clés ou tags, est une manière commune d'organiser le contenu pour l'avenir la navigation, du filtrage ou de recherche. Bien que l'organisation électronique contenu de cette manière n'est pas nouveau, une forme de collaboration de ce processus, qui a été donné le nom de «marquage» par ses promoteurs, est gagne en popularité sur le web.

Référentiels de documents ou de bibliothèques numériques permettent souvent de documents dans leurs collections qui seront organisées par mots-clés assignés.

Cependant, traditionnellement catégorisation ou d'indexation est soit réalisée par une autorité, comme un bibliothécaire, ou bien dérivés à partir du matériel fourni par les auteurs des documents (Rowley, 1995).

En revanche, le tagging collaboration est la pratique de permettre à certains - surtout les consommateurs - de joindre librement mots-clés ou des balises au contenu. Tagging collaboratif est le plus utile quand il n'y a personne dans le bibliothécaire "rôle" ou s'il y a tout simplement trop beaucoup de contenu pour une seule autorité pour faire classer; ces deux traits sont vraies pour le Web, où le marquage de collaboration ont progressé populaire.

Ce genre de tagging collaboratif offre une alternative intéressante aux efforts actuels visant à ontologies Web sémantique (Shirky 2005), qui ont été l'objet de recherches par un certain nombre de groupes (par exemple, Doan, Madhavan, Domingos Halevy 2002). Un certain nombre de maintenant, sites web très fréquentés fonctionnalité de tagging collaboratif. En règle générale, ces sites permettent aux utilisateurs de marquer publiquement et partager du contenu, de sorte qu'ils ne peuvent pas classer uniquement des informations pour eux-mêmes, ils peuvent parcourir les informations classées par autres. Il est donc tout à la fois personnelle et publique aspects des systèmes tagging collaboratif. Dans certains sites, tagging collaboratif est également connu comme "folksonomie", abréviation de "Taxonomie populaire:" Toutefois, certains se demandent si ce terme est précise (Mathes 2004), et ainsi nous éviter de l'utiliser ici. Del.icio.us, le site sur lequel nous avons effectué notre analyse, permet pour le tagging collaboratif des signets partagés site.

Yahoo MyWeb le fait aussi, et citeulike et connotea faire la même chose pour les références à des publications académiques. Quelques services permettent aux utilisateurs de balises, mais seulement le contenu dont ils sont propriétaires, pour Par exemple, Flickr pour les photos et Technorati pour les postes de weblog. Bien que ces deux sites ne sont pas, à strictement parler, le soutien tagging collaboratif, nous les citer pour illustrer la croissance de tagging dans une variété de médias .

2.2. Les tags :

Dans les applications contentes et centrales, les utilisateurs naviguent typiquement sur le contenu par les catégories ou les menus authored par les rédacteurs de site. Chaque catégorie peut avoir un certain nombre a niché les sous-catégories, en permettant à l'utilisateur de forer en bas l'arbre de sous-catégorie et trouver le contenu de l'intérêt. D'un point de vue d'expérience de l'utilisateur, une telle navigation peut être ennuyeuse. Un utilisateur pourrait devoir naviguer à travers les sous-thèmes multiples avant de trouver le bon article. Cette approche de manuellement classer des articles par catégories peut être chère et difficile maintenir sur la longue course en raison de la main-d'oeuvre impliquée, d'autant plus que le quantité d'augmentations contentes. Comme les utilisateurs produisent le contenu sur votre site, ce sera cher et quelquefois financièrement impraticable à manuellement classer le contenu par catégories être créé. Imaginez-le un site comme Flickr avec des millions de photographies et de l'effort serait exigé si vous avez essayé de manuellement classer chaque photo par catégories peut apprendre de chaque utilisateur - quelle sorte de contenu elle s'intéresse à - et dynamiquement construisez *des liens de navigation*, des liens hypertexte à d'autres articles pertinents, dont le texte ou les expressions sont aussi familier à l'utilisateur. De plus, un tel système peut être construit dans une manière efficace du prix sans devoir compter sur les rédacteurs professionnels pour classer des articles par catégories.

Les utilisateurs tagging des mots clé ajoutant les articles ou des expressions à sont articles maintenant omniprésents sur le web. Ce processus simple d'un utilisateur ajoutant des étiquettes ou des tags aux articles, en marquant les articles, en partageant des articles, ou en voyant simplement des articles fournissent dataset riche qui peut traduire dans l'intelligence, tant pour l'utilisateur que pour les articles. Cette intelligence peut être dans la forme de trouver des articles rattachés à celui étiqueté; le branchement avec d'autres utilisateurs qui ont de la même façon articles étiquetés; ou le dessin de l'utilisateur pour découvrir des tags alternantes qui ont été associées avec un article d'intérêt et au cours de cette conclusion d'autres articles apparentés .

2.2.1. La recherche par tags :

Nous l'avons vu, avec la *Folksonomie*, l'internaute peut désormais mettre en place et partager sa propre classification via les tags. L'utilisateur affecte l'étiquette de son choix à ses favoris, images, vidéos. L'avantage est indéniable : c'est l'internaute qui définit son étiquette et non plus un organisme. Mais les inconvénients sont nombreux : difficile gestion des singuliers/pluriels, problème de synonymes, de signification (que désigne cet internaute par ce tag), d'homonymes, de polysémie, d'hétérogénéité : plusieurs tags peuvent désigner les mêmes documents.

Toutefois, l'approche par tag facilite la recherche, d'une part en ajoutant une couche sémantique aux documents web et d'autre part, en permettant une navigation par concepts, de plus en plus abandonnée compte tenu de la désertion progressive des catégories des grands répertoires web par les internautes.

De plus, on peut tagger pratiquement tout type de contenu web. Ainsi, les sites *43 Things* et son équivalent français *36 Trucs* 80 proposent aux internautes de lister leurs objectifs, expériences ou projets. Chacun de ces projets est *taggé* via un ou plusieurs mots clés

2.2.2. Partageons les tags :

Oui, nous convenons, des tags sommes frais. Je suis un grand ventilateur d'intelligence collective, et ai personnellement éprouvé l'alimentation électrique de tagging collaboration. Avec mes collaborateurs chez RealTravel, nous avons construit un produit du « Web 2.0 » qui a la gestion de réseau, et tagging contents et sociaux contribués par utilisateur. Je voudrais joindre des forces avec mes collègues dans la communauté de tagging (TagCamp, 2005) pour aider à établir l'infrastructure qui permettra à des systèmes comme RealTravel d'interopérer dans un écosystème des points d'émission, services, agents, et outils qui combinent et ajoutent la valeur à l'étiquetage fait par tous ces utilisateurs. Comment faisons-nous ceci ? Vous avez deviné : créer une ontologie pour folksonomy.

Commençons par clarifier nos buts. Après tous, c'est un effort de conception d'ingénierie pas un exercice en classant le contenu du monde. Considérer deux cas d'utilisation.

Cas d'utilisation 1 : tagging collaboration à travers des applications multiples

Aujourd'hui nous pouvons tagging nos photos sur Flickr et utiliser des tags pour bookmarking dans Del.icio.us. Nous pouvons des blogs de consultation sur Technorati par des tags. Je veux tagging le contenu que je trouve sur n'importe quelle application, et je veux l'avantage de d'autres tags à travers ces applications. Ceci signifie qu'il doit y avoir une certaine voie des motifs au sujet de l'équivalence ou un rapport parmi tagging des données à travers des applications.

Par exemple, disons-moi souvent un blog au sujet de mon déclenchement à Bali sur mon site préféré de voyage. Je tagging avec les tags que qui classent mon déclenchement pour le voyage (plongée » etc. de « culture » de « aventure » «) et le site de voyage peut tagging automatiquement avec des choses comme les places j'ai visitées (« Bali » « Ubud » etc.) et mon nom d'écran. Si j'utilise des photos de Flickr dans ce blog de voyage, je veux afficher les tags existantes de Flickr quand ces photos sont affichées dans le cadre du blog de voyage (plage » etc. de « lotusflower » «). Je veux mettre les tags voyager-connexes sur les photos de Flickr, ainsi l'assistance de Flickr peut automatiquement savoir que la fleur de lotus était dans Ubud, et vice versa, ainsi l'assistance de site de voyage peut trouver des images des plages dans Bali. Quand mon blog est syndiqué par le réseau, je le veux les forces de jointure avec d'autres blogs dans Technorati ou Del.icio.us ou d'autres agrégateurs, où le monde entier peut obtenir les flots focalisés de frais, authentique, contenu utilisateur-contribué au sujet des places qu'ils veulent visiter et des choses ils veulent faire. tagging à travers de diverses et diverses applications, existantes et être créé, exige que nous permettons pour permuter, compare, et raison au sujet des données d'étiquette sans n'importe quelle une application possédant le « espace tags » ou folksonomy.

Cas d'utilisation 2 : Filtrage de collaboration basé sur tagging

Google est grand, mais il prend le travail triant par tout le bruit. Pourquoi est-ce que je devrais devoir répéter l'effort si d'autres sont déjà passés par lui ? Quand j'écris le terme « folksonomy » dans un Search Engine je vouloir pouvoir voir les résultats qu'un bon nombre de d'autres ont tagging avec folksonomy. Je veux trouver l'écriture de Shirk et de Vander Wal tout de suite parce que d'autres personnes les ont implicitement marqués au besoin lisant. Si un certain spammer a essayé de détourner un tag populaire, je veux que les masses le

rejettent avec leur tagging. Je veux être dit, sans savoir à demander, que corrélations les « folksonomy » de terme avec le terme « tagging », parce que les gens ont tagging les mêmes choses avec les deux tags. Je veux aller à Rojo et voir quels blogs tagging avec ce mot sont les plus lus, à travers tous les systèmes importants de bloguant. Dans le meilleur des cas, je veux voir using ce que mes collègues autour du monde ont tagging en soi, quelque système de tagging ils choisissent. Plus généralement, quand j'effectue le travail de la connaissance sur le Web, je veux tirer profit de tout l'autre travail que d'autres personnes ont effectué. Je veux découvrir d'autres personnes effectuant le même travail, peut-être pour partager ou se connecter vers le haut. C'est la vision qui a lancé le Web, et il pilote le but d'accélérer la connaissance et la compréhension humaines. Que prend-il ? De nouveau, ce cas d'utilisation exige qu'il y ait une conceptualisation commune de ce qui tagging des moyens et au moins une certaine voie pour qu'un service corrèle ou de connecte des données des tags d'une application à l'autre.

Comment procéder ? Je doute qu'il y ait jamais une voie simple et normalisée de rassembler, interpréter, ou utiliser des données d'étiquette. Mais nous pouvons construire le substrat pour un écosystème d'étiqueter que cela la volonté nous laisse innover et travailler vers la vision d'un tagosphere ouvert. J'argue du fait que l'ontologie est noyau à cet effet. Nous identifions une conceptualisation commune, et établissons un cahier des charges au niveau sémantique. Nous identifions et construisons les systèmes qui commettent aux caractéristiques à de divers niveaux d'engagement, et la connexion l'écosystème. En particulier, nous proposons une conceptualisation d'étiqueter cela permet l'alimentation électrique que nous voulons tout en permettant l'innovation dans la mise en place, l'optimisation, et l'extension. Nous hachons dehors ces concepts qui sont clairs, et essayons de faire des définitions non ambiguës pour des limites. Nous identifions ces concepts qui sont vagues, et nous mettons à les clarifier. Et nous présentons un cadre conceptuel pour identifier zones où les systèmes différeront. Les ontologies sont autant au sujet des motifs au sujet des incompatibilités qu'au sujet de trouver des vulgarisations.

Une ontologie tag - certains considérations de conception Avec cette vision et ces tris des cas d'utilisation à l'esprit, un groupe de personnes de la communauté de tagging commencent à travailler à une ontologie commune pour étiqueter - le TagOntology (Gruber, 2005). (Note : ce n'est pas au sujet de développer un folksonomy commun - un ensemble

commun de mots à utiliser en étiquetant. Par exemple, l'ontologie n'inclura pas des limites pour les documents de écriture de labels sous des sujets de la science ou des affaires ; elle ne sera pas pour modeler des domaines particuliers tels que la géographie ou la photographie.) Le TagOntology est au sujet d'identifier et de formaliser une conceptualisation de l'activité de l'étiquetage, et la technologie de la construction qui commet à l'ontologie au niveau sémantique. La communauté travaille également à permettre l'infrastructure aux niveaux des formats, des modèles de données, et des api. L'approche plus grande est de créer une pile logique de conception à la mise en place qui stimule l'innovation à tous les niveaux.

Concentrons ici sur la couche d'ontologie. Si développer des ontologies est comme la conception d'ingénierie (Gruber, 1995), ce qui sont certains des problèmes de conception se posant à nous ? J'offrirai une saveur pour les issues ici, et offre une certaine analyse préliminaire. Le travail réel pour résoudre des solutions est de collaboration et continu.

2.2.3. Tags structurables :

Pour faciliter l'organisation de leurs tags et pour permettre des inférences par les systèmes informatiques, SemanticScuttle offre aux utilisateurs la possibilité de créer des relations entre tags. Il existe deux types de relations : la relation d'inclusion symbolisée par le caractère « > » et la relation de synonymie symbolisée par le caractère « = ».



Figure 2.2 : Exemple de ressource vidéo employée par les enseignants

2.2.4 Le concept de noyau : tagging

Pour permettre les cas d'utilisation décrits ci-dessus, l'idée de noyau de tagging doit expliquer le plein environnement de tagging social. Du point de vue de l'utilisateur, tagging est une activité dans laquelle vous étiquetez un certain contenu que vous créez ou expérience avec une ou plusieurs tags, ou des tags. Tellement un pourrait être tenté pour le formaliser comme relation de deux-place

tagging (objet, tag)

C'est très bien si vous vivez dans un monde fermé. Mais pour permettre le filtrage de collaboration, vous avez besoin d'une certaine notion de tagger - la personne ou l'agent faisant tagging. Ainsi nous devons représenter le tagger dans notre relation, en tant qu'ainsi :

tagging (objet, tag, tagger)

Maintenant nous devons penser à la façon dont ces données pourraient être partagées. Vous ne pouvez pas laisser ceci implicite au niveau d'inter-application. Par exemple, si deux applications modelaient leurs données tag using la relation de trois-place, quand elles ont mis en commun ou ont permuté leurs données, il pourraient ressembler à ceci :

Tagging (Object1, tag1, tagger1) de // par le système 1

Tagging (Object1, tag2, tagger1) de // par le système 1

Tagging (Object1, tag1, tagger2) de // par le système 1

Tagging (Object1, tag3, tagger3) de // par le système 2

Tagging (Object2, tag1, tagger4) de // par le système 2

là où les trois premiers faits sont du premier système et du reste être de l'autre système. Si nous devons comparer des données des autres systèmes, nous ne pouvons pas supposer qu'elles toutes ont exactement les mêmes ensembles d'objets, des tags, et de taggers. Ainsi, nous devons faire explicite une certaine notion de la source, que vous pouvez penser à comme portée des espaces de noms ou d'univers de quantification pour ces objets. (Je suis tenté pour penser à la source en termes de communauté, mais c'est une traduction spécifique à

l'application.) Tellement maintenant nous avons une relation de quatre-place, avec la source comme limite formelle :

Tagging (Object1, tag1, tagger1, source1)

Tagging (Object1, tag2, tagger1, source1)

Tagging (Object1, tag1, tagger2, source1)

Tagging (Object1, tag3, tagger3, source2)

Tagging (Object2, tag1, tagger4, source2)

Ceci nous permet de dire quelque chose au sujet d'une collection de données d'étiquette, indépendant des applications spécifiques qu'ils viennent de :

Contraintes sur le « tagging » :

Pour faire toutes les conclusions valides à partir des données fusionnées ou permutées, nous avons besoin d'un engagement ontologique à la sémantique d'utilisateurs de tagging et ses trois. D'abord, considérer le critère de la concordance interne (Gruber, 1995) pour la relation elle-même. Est d'abord la notion qu'un tagger simple « vote » avec son tag, et vous pouvez seulement voter une fois. C'est-à-dire, si $tag1 = tag2$, là n'est alors aucune différence entre les premières et deuxièmes assertions ci-dessus ; il est logiquement redondant et vous pourriez continuer à leur affirmer pour toujours sans ajouter n'importe quelle information. (Vous pouvez énoncer ce de diverses voies avec des règles ou des axiomes d'inférence, mais lui tout le montant à l'idée simple de voix.) C'est un excellent exemple de pourquoi les systèmes doivent prendre des engagements ontologiques au niveau sémantique, hormis tous les accords sur des formats (Gruber, 1993). Si un système donnait la signification différente aux assertions répétées de la relation de étiquetage, alors il serait logiquement contradictoire pour combiner leurs données.

Une deuxième qualité intrinsèque de notion à tagging est ce jeu chose-que-être-étiqueté qu'un rôle dans la signification tagging cela est différent de tag ou du tagger. Une des propositions sur la table de la discussion de TagOntology est si on peut tagging une tag.

Naturellement, on peut faire un système pour enregistrer ces tuples, mais la signification n'est pas claire sur la relation de tagging sans modification. En particulier, la relation de tagging n'est pas symétrique : vous ne pouvez pas permuter le tagger et les rôles tagging et préserver la signification d'une assertion de tagging. Clarifier ainsi la signification de tagging, nous concevons un tri différent de relation ou de famille pour « metatagging » ou celui qu'il pourrait s'appeler. Un système pourrait employer une notion tag-sur-tag pour signifier que « cette tag est un synonyme de cette tag » et un autre système pourrait avoir une notion de « cette tag représente une batterie d'autres tags ». Il n'y a aucune condition que tous les systèmes partagent les mêmes notions ; une connaissance réussie partageant l'accord exige seulement qu'elles identifient clairement les différences quand elles partagent des données.

Tagging négatif :

Considérer maintenant comment traiter le filtrage de collaboration des « mauvaises » tagging des spammers. Comment est-ce qu'une foule « vote dehors » un spammer ? Il s'avère ceci exige tagging de négatif - affirmant qu'une tag ne devrait pas s'appliquer à un objet. Queest-ce qu'un engagement minimal pour le négatif tag ? On a pu modeler l'assertion de tagging négative comme littéralement inversion : « il n'est pas vrai que tagger1 ait tagging object2 avec tag3 ». Cependant, la représentation des faits importants en tant que phrases logiques plutôt que des relations mène à tous les tris d'eau profonde de calcul. Il devient plutôt difficile de s'avérer, généralement si un étiquetage s'est produit. Il tente également d'essayer d'assigner un certain genre de poids fondé au rapport, mais ceci a les problèmes semblables dans l'essai de raisonner au sujet de tagging. Si nous pouvons nous abstenir à la tentation de faire trop, il est parfaitement raisonnable d'ajouter simplement un autre argument à la relation - un argument de polarité. Ceci nous amènerait à une relation de cinq-place :

Tagging (objet, tag, tagger, source, + ou -)

Pour donner cette signification, nous pouvons sauver la contrainte que vous obtenez seulement à une « voix », positif ou négatif. De nouveau, il y a des voies de fantaisie de dire ceci dans la logique mais je pense que l'anglais réalise un travail assez bon. Vous pouvez distinguer ce que signifierait il pour quelqu'un au « untag » quelque chose par opposition à changer leur polarité. De même, vous pouvez sauver des règles d'inférence « de logique de défaut ». Pour maintenant, nous sommes contents avec des rapports comme « si vous ne

donnez pas une polarité, il se transfère sur le positif). Bien qu'officieusement spécifique, c'est toujours un engagement ontologique au niveau sémantique : vous pouvez raisonner au sujet des regroupements des faits de la forme tagging (objet, tag, tagger, source) et tagging (objet, tag, tagger, source, polarité). Si un système utilise la version de quatre places et un autre utilise la relation de cinq places, le système de cinq-place peut impliquer que les quatre places sont équivalentes à cinq variantes de place avec « + » pour la valeur de la polarité. En raison d'une ontologie partagée, les systèmes avec tagging de négatif peuvent partager des données avec ceux qui ne supportent pas la notion, et les tiers agents qui peuvent comprendre et réconcilier les différences.

Identité d'étiquette :

En conclusion, l'ontologie a besoin des définitions formelles d'identité pour chacun de ses concepts de noyau : objet, étiquette, tagger, et source. En d'autres termes, en interprétant un ensemble d'étiqueter des données, comment savons-nous quand deux objets, tags, taggers, ou sources sont identiques ? Le Web sémantique (et le RDF) offre une configuration commode pour enregistrer les espaces de noms using des URI. Par exemple, il n'est pas difficile d'imaginer permettre n'importe quoi avec un URI être l'objet d'une assertion de tagging. Cependant, que diriez-vous des tags ? Distingue les majuscules et minuscules dans les noms ? L'espace blanc ? Ces issues ou sont-ils juste détails d'implémentation de niveau sémantiques ?

Il est clair que les différents systèmes de tagging traitent aujourd'hui la puissance d'entrée, la sortie, et appairer des expressions tag différemment. Cependant, nous croyons qu'il est possible de formaliser une conceptualisation qui factorise dehors ces différences clairement, de sorte que les tiers agents puissent raisonner au sujet des différences. Une technique est de représenter une fonction des noms aux étiquettes. Par exemple

$f(\text{« San Francisco »}) = \text{tag1}$

$f(\text{« San Francisco »}) = \text{tag2}$

$f(\text{« San Francisco »}) = \text{tag3}$

Alors on peut sauver les axiomes clairs qui définissent comment un système particulier traite apparié nommé. L'on a pourrait indiquer que $\text{tag1} = \text{tag2}$, des autres qui $\text{tag2} = \text{tag3}$, et ainsi de suite. Il est également possible de modéliser la fonction l'autre voie, celle qu'une tag donnée a un cname nommé canonique (tag) = " chaîne de caractères ". Alors des différences parmi les formes extérieures des tags sont liées dans l'application (c.-à-d., ne pas s'inquiéter comment des tags sont entrées et affichées dans l'application, mais insister sur le fait que n'importe quelle exportation des données tag à d'autres systèmes utilise seulement le nom canonique).

Les issues semblables surgissent-elles pour la portée des identificateurs pour des taggers - devraient-elles être relativement à la source ou être exigées pour scoped universellement par quelque chose comme un URI ? Si elles scoped par URI, ont-elles également des noms de chaîne extérieurs (« noms d'écran ») qui peuvent être appariés par de tiers systèmes, et par lesquels règles d'identité ? La source représente-t-elle un ensemble de taggers ou d'autre chose ? Y a il n'importe quelle portabilité d'identité à travers des applications, et si oui, par quel mécanisme (enregistrement central, allumette coïncidente de chaîne de caractères, FOAF (Miller et Brickley, 2005), etc.) incompatibilités comme au sujet de trouver des vulgarisations.

2.2.5. Pourquoi les utilisateurs étiquettent-ils ? :

Dans sa forme la plus fondamentale, les gens peuvent tag des choses pour organiser des articles d'intérêt et souvenez-vous d'eux. Par exemple, si vous avez un grand nombre de dossiers sur votre ordinateur, vous créez normalement des chemises pour organiser vos dossiers. Le fait de tagging est semblable - vous créez des catégories ou les seaux et les associez à l'article d'intérêt. Un article peut être tagging avec les tags multiples, comme montré dans la figure 8, un screenshot d'Amazon.com, où on permet des utilisateurs d'ajouter des tags multiples à un article. En tagging des articles, les utilisateurs peuvent utiliser des tags cela leur a du sens et ne doit pas utiliser la classification du propriétaire content ou site.

Les utilisateurs tagging aussi des articles pour qu'ils puissent partager les renseignements avec d'autres, trouver rattaché les articles que d'autres ont tagged dans la même catégorie et aussi quand ils veulent être trouvé par d'autres (principalement dans les sites interconnectant socialement).

Comme montré dans la figure 6, supposons qu'un utilisateur s'intéresse à la direction et a placé une tag de direction pour ses articles d'intérêt. Pour la durée du temps, si l'utilisateur voudrait voir tous les articles rattachés à la direction, il peut regarder tous les articles qui ont été tagged avec la direction. Les tags aident des utilisateurs à organiser et trouver des articles d'intérêt.



Figure 2.3 : L'Amazone permet aux utilisateurs de tags un produit et voir comment d'autres ont tagging le même produit

Un autre trait c'est activement utilisé, surtout pour les applications de mise en réseau sociales, est permettre aux utilisateurs d'explorer des tags . Dans notre exemple, n'importe quel utilisateur serait capable de cliquer le le lien de direction et voit d'autres articles qui ont été étiquetés par d'autres utilisateurs avec le même tag. Cela les aide à découvrir de nouveaux articles d'intérêt. L'utilisateur peut regarder aussi que les utilisateurs ont utilisé la même tag, en les aidant ainsi à trouver d'autres utilisateurs avec les intérêts semblables.

Vous pouvez trouver aussi des tags semblables. Le même article peut avoir été tagged par d'autres avec différents mots clé. Par suivant des liens, vous pouvez associer des tags semblables. Par exemple, dans la figure 8, le même article a été tagged que les réseaux sociaux, la sociologie, la direction, marketing, et cetera. Ces quatre tags peuvent être semblables, surtout s'ils sont répétés systématiquement à travers les articles multiples.

2.3. Folksonomie :

On parle beaucoup à l'heure actuelle, dans le cadre des projets de Web 2.0, des systèmes de "folksonomie", permettant aux utilisateurs d'un outil de voter, de donner leur opinion sur tel ou tel site ou page web, ces indications permettant d'améliorer notamment les résultats de recherche des futurs utilisateurs.

Le mot "folksonomie" est apparu en 2002 sur Internet pour décrire une expérience de classement de ressources effectué librement et spontanément par des utilisateurs du réseau. C'est l'informaticien Thomas Vander Wal¹ qui est à l'origine de ce terme, utilisé depuis dans le concept de Web 2.0. Plus largement, la folksonomie désigne les actions de classification, annotation et indexation de tous types de documents Internet réalisées de façon naturelle et massive par les utilisateurs du réseau. L'objectif est de bénéficier des efforts de tous pour trouver plus facilement et plus vite des ressources de bonne qualité.

2.3.1. Définition et présentation de la folksonomie :

Une folksonomie est un système de catégorisation qui émerge de tagging collaboratif sur Internet. Il s'agit donc simplement d'un ensemble de mots que les membres d'une même communauté informelle (e.g., les utilisateurs de Flickr) utilisent pour étiqueter des contenus en ligne. Une folksonomie est donc liée à un site communautaire bien particulier : par exemple, la folksonomie de Flickr est différente de celle de Dailymotion. Ce néologisme est la Francisation du terme *folksonomy*, qui provient lui-même de la contraction des termes *folks* (gens) et *taxonomy* (bien qu'il ne s'agisse pas d'une taxonomie) qui évoque l'idée d'une catégorisation par les gens.

Chaque participant catégorise l'information qui l'intéresse selon son propre point de vue et accepte de partager son classement avec les autres utilisateurs. L'objectif est de bénéficier des efforts de tous pour chercher mieux et plus vite, de faire des découvertes en rapport avec

¹**Thomas Vander Wal** est l'auteur du néologisme *folksonomy* traduit en français par folksonomy. Il est consultant de la société InfoCloud Solutions fondée en 2005 à partir de ses travaux

ses centres d'intérêts. La folksonomie industrialise en quelque sorte le principe de bouche à oreille que nous pratiquons tous dans la vie courante.

En plus du terme folksonomie, on rencontre dans Internet plusieurs autres équivalents pour désigner ce concept en français : *socionomie* (*socio-*, de *social*, *société*, ce mot aurait un sens différent en économie), *peuplonomie* (*peuplo-*, préfixe français inexistant; mot non transparent), *populonomie* (de *populo*, *populace* « bas peuple », connotation péjorative), *personomie* (de *personne*, manque de précision) et *potonomie* (de *pote* « ami », mot familier, non pertinent dans le contexte). Cependant, ces derniers, très peu attestés, n'ont pas été retenus, car ils constituent, à notre avis, des tentatives de francisation peu satisfaisantes sur le plan terminologique et viennent concurrencer inutilement le terme folksonomie, déjà largement utilisé.

2.3.2. Folksonomy comme données qui sont émergentes de l'information partagée :

Il n'y a pas très longtemps, les observateurs vifs de l'Internet (Vander Wal, 2004), (Sterling, 2005), (de Mieszkowski, 2005) et les inventeurs du logiciel social (Shachter, 2003), (article truqué et Butterfield, 2003) ont commencé à noter que ces personnes qui ne savent pas les programmes informatiques heureusement « tagging » avec des mots-clés le contenu qu'ils ont créé ou ont rencontré. Naturellement, tagging de mot-clé est rien de neuf ; l'observation intéressante est que quand ces gens font leur tagging dans un espace public, la collection de leurs associations de mot-clé/valeur devient une source utile de données dans l'agrégat. Par conséquent le terme « folksonomy » - l'écriture de labels émergente d'un bon nombre de choses par des personnes dans un contexte social. Thomas Vander Wal, qui est crédité de la limite, souligne que résulter folksonomy n'est pas une taxonomie ou même une catégorisation de collaboration (Vander Wal, 2004). Au moins c'était l'observation et l'intention initiales pour la limite. Aujourd'hui, l'étiquetage est un phénomène répandu popularisé par des applications telles que le social bookmarking (Del.icio.us) et partageant social de photo (Flickr). Dans ces applications, les données émergentes des actions des millions de gens ordinaires et non formés faisant des choses pour leurs propres intérêts locaux sont plutôt utiles. Pour bookmarking, tagging aide à parer le bruit Spam-induit dans des moteurs de recherche, et pour la photo le partage, étiquetant donne à ceux les moteurs de recherche basés par texte une chance de combat.

2.3.3. Traitements sur les folksonomies :

Une des limitations couramment reconnue aux folksonomies (Mathes, 2004) est la variabilité d'écriture des tags supposés équivalents comme "écologie" et "ecologie". Une solution possible pour traiter ce problème consiste à mesurer la distance d'édition entre les tags (par exemple de type distance de Levenshtein (1966)), et au delà d'un certain seuil, de considérer ces tags équivalents. Specia & Motta (2007) ont appliqué cette méthode sur un extrait de la folksonomie de delicious.com, et exploité des bases de connaissances externes (Wordnet) et quelques règles simples pour sélectionner le libellé du tag le plus représentatif de ses variantes orthographiques.

Un autre type de traitement des folksonomies consiste à mesurer la distance de "similarité" entre tags en se basant sur les liens entre les tags, les ressources, et les utilisateurs (Mika, 2005). (Cattuto et al., 2008) distinguent sur ce point différents types de mesures de similarité : les mesures basées sur une fréquence de cooccurrence "simple" de deux tags sur une même ressource, ou les mesures distributionnelles, qui prennent en compte trois types de contextes d'association des tags. Chaque contexte correspond à un espace vectoriel prélevé dans l'espace vectoriel global de la folksonomie.

Les mesures distributionnelles prennent en compte l'association des tags : (1) via leur usage par un même utilisateur (contexte utilisateur-tag), ou (2) via leur usage pour une même ressource (contexte ressource-tag), ou (3) via leur associations communes avec d'autres tags (contexte tag-tag).

Afin de caractériser en termes sémantiques ces différentes mesures de similarités entre tags, Cattuto et al. proposent d'exploiter la structure hiérarchique de Wordnet (Fellbaum, 1998) pour les tags dont le libellé est présent dans cette base lexicale. L'issue de cette expérience montre que les tags associés via des mesures de cooccurrences simples tendent à entretenir des relations de subsomption, alors que les tags associés via une mesure distributionnelle de similarité dans le contexte tag-tag tendent à se situer au même niveau hiérarchique, soit partageant le même parent, soit le même grand parent.

Cattuto et al. expliquent que l'association des tags via leur cooccurrence sur une même ressource renvoie à leur utilisation simultanée dans le même acte de tagging ou l'utilisateur a

tendance à couvrir différents niveaux de généralité. Par exemple, les tags “java” et “programming”, ou encore “tobuy” et “shopping” sont fréquemment utilisés IC 2009 simultanément, et on peut supposer que, du point de vue du “tagueur”, ces tags ont des niveaux différents de généralité. Le lien mesuré par la mesure distributionnelle dans le contexte tag-tag associe des tags ayant des schémas de cooccurrence similaires mais qui ne sont que peu ou pas utilisés simultanément. Ce cas de figure correspond par exemple aux tags “tobuy” et “whishlist” qui ne sont pas utilisés simultanément mais plutôt conjointement avec le tag “shopping”.

La principale limite à l'utilisation de Wordnet comme base de connaissance est que cette ressource termino-ontologique inclue peu de termes spécifiques à un domaine, alors qu'ils sont fréquents dans les folksonomies. Des ressources plus spécifiques à un domaine pourrait donc permettre d'élargir la portée de la validation sémantique des liens de similarités entre certains tags. Cependant la rareté de telles ressources, et la limite de leur couverture d'un domaine repousse toujours plus loin le problème. L'expertise des utilisateurs d'un système semble en définitive la plus adaptée, mais aussi la plus complexe à exploiter si on cherche autant que possible, afin d'éviter toute surcharge cognitive, à limiter l'effort de contribution nécessaire à la formalisation de cette expertise.

2.4. Tagging et Taxonomie :

Les partisans de de tagging collaboration, typiquement dans la communauté web logging, souvent le contraste tagging-ont basé des systèmes de taxonomies. Tandis que ce dernier sont hiérarchiques et exclusifs, les anciens sont non hiérarchiques et inclus. Taxonomies familières inclure le système de Linnaean de classier des choses vivantes, Catégorie décimale de Dewey pour des bibliothèques, et fichier électronique systèmes pour organiser les fichiers électroniques. Dans de tels systèmes, chacun l'animal, livre, classent et ainsi de suite, est dans une catégorie non ambiguë ce qui est alternativement dans pourtant le plus général. Par exemple, les lions et les tigres tombent dans le genre panthera, et les chats domestiques dans le genre felis, mais le panthera et le felis sont une partie de felidae de famille, dont les lions, les tigres et les chats domestiques sont toutes les cloisons.

De même, les livres sur la géographie de l'Afrique sont dans la catégorie de système décimal de Dewey 916 et des livres sur l'Amérique du Sud dans 918, mais tous les deux sont englobés par la catégorie 900, couvrant tous les sujets dedans géographie.

En revanche, tagging n'est ni exclusif ni hiérarchique et peut donc dans quelques circonstances avoir un avantage plus de taxonomies hiérarchiques. Par exemple, considérer un hypothétique chercheur qui télécharge un article concernant l'indigène d'espèce de chat en Afrique. Si le chercheur voulait organiser le tout elle téléchargerait des articles dans une hiérarchie des chemises, là sont plusieurs options hypothétiques, dont nous considérons quatre

1. c:\articles\cats tous les articles sur des chats
2. c:\articles\afrika tous les articles sur l'Afrique
3. c:\articles\afrika\cats tous les articles sur les chats africains
4. c:\articles\cats\afrika tous les articles sur des chats d'Afrique

Chaque choix reflète une décision au sujet de l'importance relative de chacun caractéristique. Les noms du dossier et les niveaux sont dans eux-mêmes instructifs, dans celui, comme des étiquettes, ils décrivent l'information retenue dans elles (Jones et autres 2005). Les chemises comme 1. et 2. Font le central le fait que les chemises sont au sujet des « chats » et de la « Afrique » respectivement, mais elide toutes les informations sur l'autre catégorie. 3. et 4. organiser les fichiers par les deux catégories, mais établir le premier aussi primaire ou plus saillant, et la seconde que secondaire ou plus spécifique. Cependant, le regard dans 3. pour un fichier dans 4. sera stérile, et ainsi le contrôle des emplacements multiples devient nécessaire.

En dépit de ces limitations, il y a plusieurs bonnes raisons d'imposer une telle hiérarchie. Bien qu'il puisse y avoir trop de chemises dans a hiérarchie, particulièrement une créée au petit bonheur, efficacement la hiérarchie des fichiers organisée bondit d'une manière ordonnée et clairement a le contenu de la chemise. À la différence d'une recherche mot--basée, où le chercheur ne peut pas être sûr qu'une requête a renvoyé tous les éléments appropriés, une hiérarchie de chemise assure le chercheur que tous les fichiers qu'elle contient sont dans une place stable.

2.5. Le tagging et le web sémantique :

Le Web sémantique (SW) s'est avéré être un outil d'intégration des données utiles, facilitant l'échange fructueux de données hétérogènes, en particulier dans les zones tels que l'e-science et la médecine. Toutefois, comme on le sait, il y a plus coûteuses-la tête dans l'utilisation de la SW, en particulier, les efforts nécessaires à la construction, et le maintien, ontologies utiles et à acquérir les riches et les RDF bien structuré peut être relativement élevés, un fait souvent accusés de ralentir l'adoption généralisée des Web sémantique. Web 2.0, et la notion de tagging de la Communauté, est prometteuse comme une alternative à moindre coût, produire des structurés modèles sémantiques à travers l'étude de la sémantique émergente. Il a fait valoir que pour pouvoir utiliser les connaissances intégrées dans folksonomies peut entraîner à la construction des ontologies peu profondes qui sont plus réceptifs aux changements des connaissances plus temps.

Néanmoins, nous ne devrions pas penser que le Web 2.0 et le SW, les étiquettes et RDF, folksonomies et ontologies sont en concurrence pour le même espace. Folksonomies sont essentiellement un développement dans la recherche d'information, une variante intéressante sur le thème de recherche par mots clés. Ceci les rend particulièrement intéressants dans le contexte des recommandations du film: ils aident à répondre à la question "comment puis-je trouver films pertinents à la notion à laquelle je m'intéresse." ontologies sont des outils de intégration de données: ce sont des tentatives de réglementer une partie du monde des données, et à mappages de faciliter et d'interactions entre les données détenues dans des formats disparates ou endroits.

La question importante à l'égard de la technologie et SW Web 2.0 n'est pas comment gérer un compromis, mais plutôt, comment les utiliser ensemble pour le meilleur avantage. Beaucoup dépendra du contexte d'utilisation particulier, mais dans le cas de recommandation film, une architecture assez basique se présente. L'utilisation du Web 2.0 Données pour l'application de la recommandation du sens, telle qu'elle ressort des tagging basés sur des perceptions. Folksonomies, étant structures organiques qui reflètent les utilisateurs ont la compréhension des ressources, peuvent fournir une meilleure base pour l'expression des intérêts de l'utilisateur. Cette idée a été étudiée dans le cadre de

sociale pour construire un système de recommandation Web en page et des résultats encourageants.

Néanmoins, l'hypothèse avec laquelle nous travaillons est de savoir si nous pouvons améliorer la performance des systèmes de recommandation, en donnant les systèmes d'accès à une plus grande quantité d'informations, ce qui implique la nécessité d'intégrer pertinentes données acquises à partir de sources hétérogènes. Ceci suggère immédiatement un rôle à jouer technologies SW. Comme indiqué, les questions qui seront abordées dans cette partie de l'architecture inclure le développement d'une ontologie appropriée et l'acquisition de RDF sans monter le coût de développement.

Conclusion :

Ce qui est plus important qu'une ontologie spécifique, cependant, est la notion plus générale que les techniques du Web sémantique, telles que le cahier des charges formel des données structurées et du raisonnement à travers des points d'émission disparates, peuvent s'appliquer au Web social. L'étiquetage des données offre une fenêtre intéressante dans l'intersection du raisonnement formel (inférence logique, analyse traitante et linguistique d'interrogation de bases de données) et des données semistruées avec la sémantique contexte-dépendante (des étiquettes et des groupements de contenu, d'identités en ligne des personnes). Les assertions d'étiquette signifient différentes choses dans différentes applications, pourtant elles ne doivent pas avoir une sémantique unifiée à être comparable à travers des sources. Le processus de développer une ontologie de données d'étiquette nous force à identifier les genres de suppositions ontologiques faites par diverse source de données d'étiquette, et à spécifier un vocabulaire pour énoncer ces suppositions. Avec une ontologie de données d'étiquette, ou les ontologies semblables pour d'autres données sociales, nous pourrions permettre des technologies pour rechercher, agréger, et connecter les personnes et le contenu qu'elles contribuent dans tout le Web. En même temps, les données riches des millions d'êtres humains actifs et participants pourraient offrir le carburant pour le développement des systèmes qui tapent l'alimentation électrique de l'intelligence collective (Engelbart, 1963).

Chapitre III

Présentation des systèmes de tagging existants

Introduction :

Le processus *tagging collaboratif*, utilisé par un certain nombre de sites internet (del.icio.us, flickr, youtube, dailymotion,...), produit une classification collective des contenus. En effet, en agrégeant automatiquement les tags utilisés par les différents utilisateurs pour classifier leurs documents ces systèmes permettent une indexation collective des documents, l'ensemble des tags constituant l'index.

En ce chapitre, nous mettons en référence de nombreux systèmes de tagging pour afficher des variations d'architecture et d'incitations. Nous n'analysons pas la plupart de ces sites potentiellement éphémères détaillés, bien que nous fournissions des références à elles afin de rectifier le lecteur avec des exemples.

Pour la lisibilité, voici une brève description des sites que nous mettons en référence. Il y a beaucoup d'autres systèmes de tagging en existence, mais nous avons choisi douze qui sont représentant de la diversité de ceux qui sont actuel bons utilisés.

Del.icio.us (<http://del.icio.us>) : « un site bookmarking social, » permettant à des utilisateurs de sauvegarder et pages Web et ressources de tags.

Flickr (<http://www.flickr.com>) : une photo partageant le système en permettant à des utilisateurs d'enregistrer et étiqueter leurs photos personnelles, comme mettre à jour un réseau des contacts et de l'étiquette d'autres des photos.

YouTube (<http://www.youtube.com>) : un système de partage visuel permettant à des utilisateurs de télécharger le contenu vidéo et de le décrire avec des étiquettes.

Yahoo! MyWeb2.0 (<http://myweb.yahoo.com>) : semblable à Del.icio.us, mais comprendre un réseau social des contacts.

CiteULike (<http://www.citeulike.org/>) : un site permettant à des utilisateurs d'étiqueter des citations et des références, par exemple papiers ou livres scolaires.

En particulier jeu (<http://www.espgame.org/>): un jeu d'Internet d'étiqueter où des utilisateurs sont aléatoirement appareillés les uns avec les autres, et essayer de deviner que les étiquettes l'autre utiliseraient une fois présentées avec a photo aléatoire.

Last.fm (<http://www.last.fm>) : une base de données de l'information de musique permettant à des membres d'étiqueter des artistes, des albums, et des chansons

Technorati (<http://www.technorati.com/>) : un agrégateur et une recherche de weblog usinent permettre à des auteurs de blog de tagging leurs poteaux.

LiveJournal (<http://www.livejournal.com/>) : un site Web de weblog et de communauté permettant à des utilisateurs d'étiqueter leur profil personnel, avec différents poteaux de blog

(<http://upcoming.org/>) : une base de données d'événements de collaboration où les utilisateurs peuvent écrire de futurs événements (par exemple, concert, des objets exposés, des jeux, etc.) et les étiqueter

En raison de leur popularité, social étiquetant les systèmes ont devenu la couverture une large gamme de ressources et de communautés, enjambant l'intervalle entier des incitations décrites dans la section précédente.

Au lieu de classifier simplement une longue liste d'outils potentiellement éphémères, nous donnerons un exemple complémentaire à ceux fournis dans les anciens travaux. Les systèmes que nous avons choisi d'étudier sont Delicious Flickr et youtube .

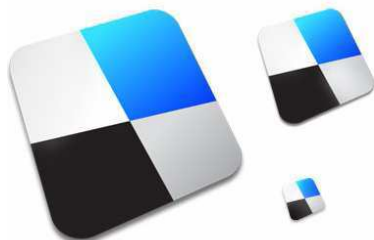
3.1. Delicious :

La recherche sur internet amène le plus souvent l'internaute à collectionner un grand nombre de pages internet qu'il souhaite mettre de côté afin de les lire plus tard.

Ces pages internet (par le biais de leur URL, c'est à dire l'adresse unique qui permet d'y accéder par le navigateur) sont le plus souvent stockées dans les "favoris" du navigateur internet utilisé. En un ou deux clics, ces pages internet sont rangées les unes après les autres sans réelle organisation si ce n'est celle de la chronologie.

Grâce à l'émergence de nouvelles applications, des outils permettent de gérer en ligne, ses fameuses sources que l'on nomme aussi favoris, signets ou encore marques-pages. En l'occurrence, cette dernière appellation est celle utilisée par un des outils les plus complets dans la gestion de marques-pages : Delicious, <http://delicious.com>, propriété de la société Yahoo!

3.1.1. A quoi sert delicious :



Delicious est un site web qui permet de sauvegarder ses marque-pages, de les classer par mots clés (aussi appelé tag, que l'on peut traduire par étiquette, mais que nous ne traduirons pas) et aussi et surtout de pouvoir les partager avec d'autres utilisateurs. Cela permet entre autre de faire des recherches dans les favoris identifiés par les utilisateurs.

Delicious est désormais l'un des outils de "bookmarking" (que l'on pourrait traduire par "marquage de pages", mais que nous traduirons "gestion de marque-pages") les plus populaires. Il s'agit d'un service entièrement gratuit et sans publicité.

Le travail collectif de recherche et de discussion en ligne au sein d'un groupe de travail, d'une association ou d'une ONG amène le plus souvent plusieurs personnes à recueillir un maximum d'informations et de pages Internet, tous susceptibles d'être partagés entre les membres d'une même équipe.

L'envoi de mails à un groupe de personne est souvent la solution la plus utilisée pour partager une information. Delicious se positionne comme un outil de centralisation et d'archivage de toutes ces recherches et permet à un nombre infini d'utilisateurs distants de pouvoir partager et nourrir leurs requêtes et leur réflexion.

Del.icio.us est un outil qui a été développé dans une optique communautaire par un diplômé de Carnegie Mellon University, Joshua Schachter, avec l'intention de pallier aux *faiblesses* des traditionnels signets à savoir : leur organisation, leur faible mobilité et leur *individualisme*.

Grâce à del.icio.us, vous pourrez (tout à fait gratuitement) organiser vos signets, vous assurer d'y avoir accès à partir de n'importe quel ordinateur branché à Internet et bénéficier du travail d'autres internautes grâce à qui vous serez *automatiquement* dirigés vers les meilleures ressources qu'Internet peut offrir sur les sujets qui vous intéressent.

Tout ça en quelques clics... et sans aucune connaissance technique particulière!

L'interface du site repose sur du HTML simple, ce qui rend le site facile d'utilisation. del.icio.us propose également de syndiquer son contenu par RSS et repose sur la technologie des tags. Les tags, sous la forme d'un mot (par exemple : sports, cinéma, Internet, etc) permettent de retrouver facilement les différents sites ayant un rapport avec le mot du tag. Les tags sont choisis par l'utilisateur lui-même, ce qui lui permet de gérer entièrement ses marque-pages.

Nous allons maintenant parler de la dynamiquement de delicious

3.1.2. dynamique Delicious :

Del.icio.us, ou Delicious, est un tagging de collaboration système de signets web que son créateur, Joshua Schachter, appelle un «gestionnaire de signets sociaux" [16]. Nous avons analysé les données provenant de Delicious pour découvrir des modes parmi les utilisateurs, les tags et les URL. Nous décrivons brièvement Delicious et d'analyser les tags dans cette section, et d'analyser signets et des URL dans la section suivante. Enfin, nous discutons la valeur de cette collaboration a généré données.

Une grande partie de la même manière les utilisateurs à gagner dans les signets leur navigateur, ils peuvent enregistrer des signets dans Delicious, à la place. L'avantage de le faire est que, une fois ses signets sont sur le web, ils sont accessibles depuis n'importe quel propre navigateur d'ordinateur, pas seulement de l'utilisateur. C'est utile si l'on utilise plusieurs ordinateurs, à la maison, au travail, l'école, et ainsi de suite, et est présenté comme l'un des Delicious principales caractéristiques.

Une fois les comptes utilisateurs sont créés, ils peuvent alors commencer bookmarking pages web; chaque signet dossiers la page Web URL et son titre, ainsi que le moment dont le signet est créé. Les utilisateurs peuvent également choisir au tag le signet avec plusieurs tags, ou des mots clés, des leur choix. Chaque utilisateur a une page personnelle sur laquelle leurs signets sont affichés, cette page se trouve à <http://del.icio.us/username>. Sur cette page, tous les signets l'utilisateur n'a jamais créés sont affichées dans le sens inverse l'ordre chronologique avec une liste de toutes les tags du l'utilisateur n'a jamais donné à un signet. En sélectionnant un mot-clé, on peut filtrer liste de signets ses signets de sorte que seulement avec cette balise sont affichés.

Delicious est considéré comme «social» parce que, non seulement peut on voir ses propres signets, on peut également voir tous tous les autres utilisateurs de signets. La première page du Delicious plusieurs spectacles de la plus récemment ajoutées signets, y compris les tags qui leur est donné, qui les a créés, et combien d'autres personnes ont que signet en commun. Il ya aussi un «populaire» page, qui montre les mêmes informations pour les URL sont actuellement les plus populaires. On peut aussi voir tout page personnelle d'autres utilisateurs et de la filtrer par tag, beaucoup de la façon dont on peut soi-même. Grâce aux pages des autres personnels et le «populaire» page, les utilisateurs peuvent avoir une idée de ce que les autres à trouver intéressante. En naviguant sur des personnes en particulier et les étiquettes, les utilisateurs peut trouver des sites qui présentent un intérêt pour eux et peut trouver des gens qui ont des intérêts communs. Cela aussi est présentée comme une des principales caractéristiques de Delicious.

Ces deux caractéristiques - le stockage des signets personnels et le caractère public de ces signets - sont quelque peu en contradiction avec l'autre. Les données que nous présentons ci-dessous confirme que les utilisateurs signet principalement pour leur propre

bénéfice, et non pour le bien collectif, mais peuvent néanmoins constituer une bonne utilité publique .

3.1.2.1. Les données :

Notre analyse a été réalisée sur deux ensembles de Delicious données, ce qui nous avons récupéré entre le matin du Friday, Juin 23 et la matinée du lundi Juin 27,2005 [16] . Les deux ensembles de données ont été créés d'abord récupérer RSS flux publics et puis l'exploration d'une partie du site Web.

La première série («populaire») contient toutes les URL qui est apparu sur la délicieuse «populaire» au cours de cette page délai. Notre ensemble de données contient tous les signets jamais affecté à chacune de ces URL indépendamment du temps, de sorte que pour chaque URL nos données contient son histoire complète au sein du système. Un total de 212 URLs et 19.422 signets constituent cet ensemble de données.

Notre deuxième série de données («peuple») se compose d'un hasard échantillon de 229 utilisateurs qui ont publié au cours de la Delicious délais ci-dessus. Notre ensemble de données contient tous les signets jamais enregistré par les utilisateurs, indépendamment de l'heure, de sorte que pour chaque utilisateur de notre jeu de données contient que l'utilisateur complète histoire. Un total de 68.668 signets composent ce ensemble de données.

Nous commençons par examiner l'utilisation de la balise individuelle utilisateurs. Comme les utilisateurs signet nouvelles URL, ils créent des balises pour les décrire. Au fil du temps, les listes des utilisateurs de balises peuvent être considéré comme descriptif des intérêts qu'ils détiennent ainsi que de leur méthode de classement de ces intérêts.

Tout d'abord, nous regardons l'activité des utilisateurs à l'égard de leurs utiliser la balise. Ensuite, nous examinons plus en eux-mêmes tags détail.

3.1.2.2. Activité de l'utilisateur et la quantité tag :

Comme on pouvait s'y attendre, les utilisateurs sont très variables de la fréquence et la nature de leur utilisation Delicious. Dans notre «Ensemble de données de personnes, il n'y a qu'une faible relation entre l'âge de compte de l'utilisateur (c.-à-temps depuis qu'ils ont créé le

compte) et le nombre de jours sur lequel ils ont créé au moins un signet ($n = 229$; $R^2 = 0,52$). C'est, certains utilisateurs d'utiliser Delicious très fréquemment, et d'autres moins souvent. Notez que ces données ne comprennent pas les utilisateurs qui avaient déjà utilisé Delicious mais il s'est arrêté, car ils étaient tous des utilisateurs actifs au moment où l'ensemble de données ont été recueillis.

Plus intéressant, il n'y a pas une relation forte entre le nombre de signets d'un utilisateur a créé et le nombre de tags qu'ils ont utilisés dans les signets ($N = 229$; $R^2 = 0,33$). La relation est faible au bas de l'échelle, les utilisateurs de moins de 30 signets ($N = 39$; $R^2 = 0,33$), et encore plus faible à l'extrémité supérieure, les utilisateurs de plus de 500 signets ($n = 36$; $R^2 = 0,14$). Certains utilisateurs ont relativement importante séries de tags, et les autres utilisateurs ont relativement peu ensembles (figures 3.1 et 3.2).

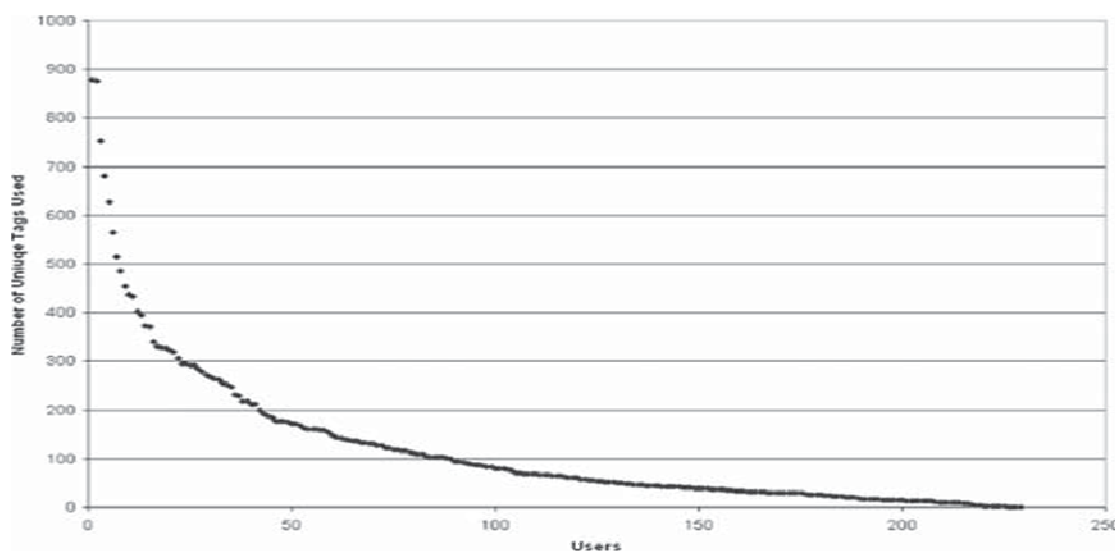


Figure 3.1 : Le nombre de tags dans la liste de tag de chaque utilisateur, en par ordre décroissante

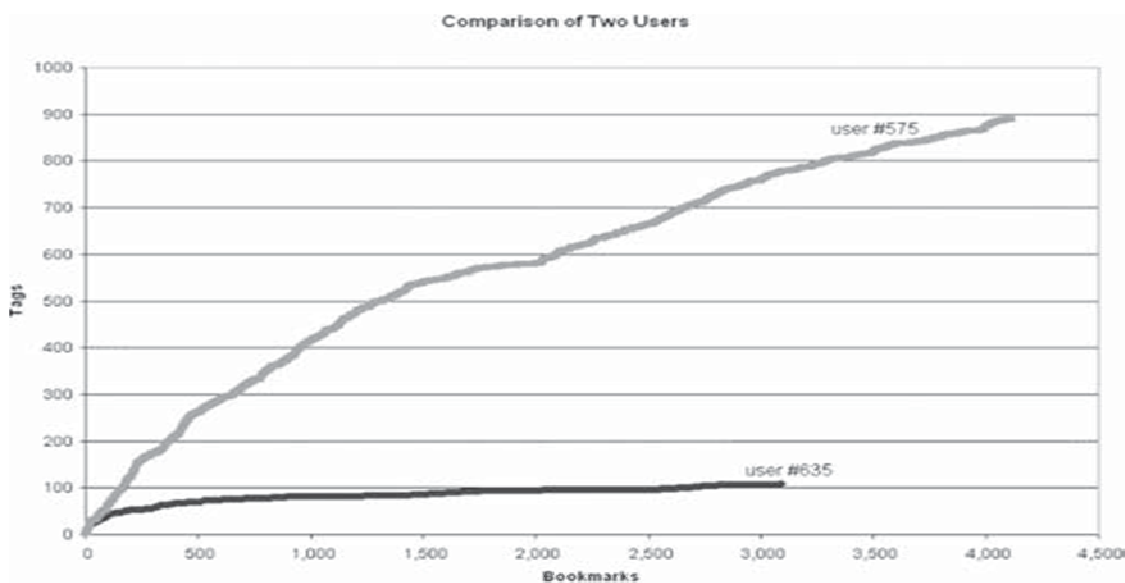


Figure 3.2 : Deux utilisateurs extrêmes »(# 575 et # 635) la croissance tag. Comme Ils ajoutent plus de signets, le nombre de tags qu'ils utilisent augmente, mais à des taux très différents.

Listes tag utilisateurs croître au fil du temps, à mesure qu'ils découvrent de nouvelles tags intérêts et ajouter de nouveaux à classer et à décrire eux. Tags peuvent présenter des taux de croissance très différents, Toutefois, ce qui reflète la façon dont les intérêts des utilisateurs de développer et changer avec le temps.

Figure 3.3 A et B montre comment l'utilisation des chaque augmentation tag que chaque utilisateur ajoute plusieurs signets au fil du temps. Pour chaque utilisateur, deux des usages de ces balises «grandir régulièrement, reflétant les intérêts continue a tagué d'une manière cohérente manière. Un tag se développe rapidement, reflétant une nouvel intérêt ou d'un changement dans le marquage pratique. Il est possible que la balise nouveau en pleine croissance représente un nouveau d'intérêt ou d'une catégorie à l'utilisateur. Une autre possibilité est que l'utilisateur a choisi d'établir une nouvelle distinction parmi leurs favoris, ce qui peut s'avérer problématique pour l'utilisateur.

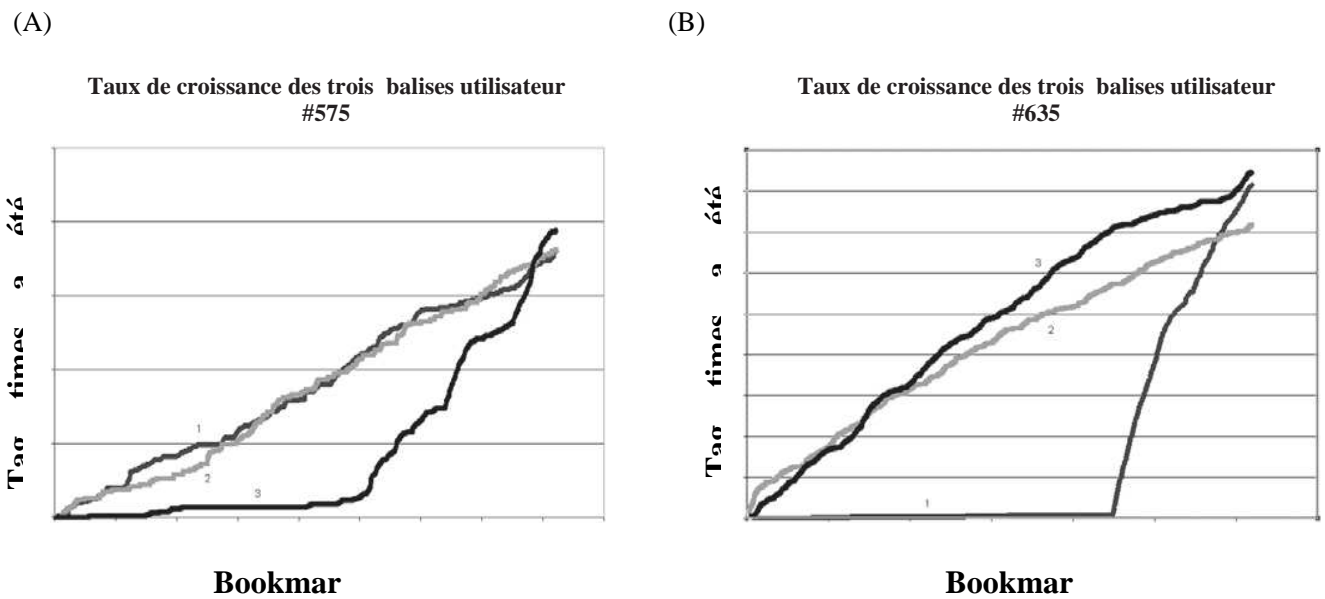


Figure 3.3 : 3Taux de croissance des trois balises sélectionné pour deux utilisateurs (#575 et 635)

Parce que sensemaking est un processus rétrospectif, informations doivent être respectées avant que l'on peut établir son sens [39]. Par conséquent, une distinction peut passer inaperçue pendant une longue période jusqu'à ce qu'il soit finalement créé par la individuels, qui continue alors de trouver cette distinction important pour donner un sens de l'information à venir. Depuis trouver de l'information est déjà rencontrés extrêmement important [43], cette situation est très problématique pour les informations passées. Par exemple, l'utilisateur # 635 (figure 3.3 b) ne pas utiliser de balise '1' jusqu'à environ la 2200e

Ajouter à mes favoris. Si «tag 1 'constitue en effet une nouvelle distinction parmi les types d'éléments de ce signets utilisateur, si Delicious ne permettent aux utilisateurs de modifier précédente signets, il serait difficile de revenir sur chacun des les 2200 plus tôt signets pour décider s'il faut ajouter «Tag 1 'pour eux. En outre, si à l'avenir de cet utilisateur a besoin pour filtrer ses signets par tag 1 ', alors aucun signet avant la 2200e seront récupérées, ce qui compromet la utilité pratique de la balise .

Ensuite, nous regardons ce qui tags fonctions de lecture dans les signets.

3.1.2.3. Types de tags :

Tagging, comme indiqué plus haut, est un acte de l'organisation par l'étiquetage, une façon de donner du sens à de nombreux discrets, des articles variés en fonction de leur sens. Par regardant ces balises, nous pouvons examiner quels types de distinctions sont importantes pour tagueurs. Il ya une certaine discussion entre les Delicious le marquage communautaire de savoir si une balise est correctement considéré comme descriptif de la chose lui-même, ou descriptif de la catégorie dans laquelle la chose diminue [16]. Cependant, nous ne voyons aucune contradiction entre ces deux types de tag. Si une catégorie est définie comme circonscrire de nombreux objets avec un particulier la propriété, nous avons naturellement examiner chacun de ces objets d'avoir cette propriété. Dans notre estimation, le champ d'application le mot-clé - si elle décrit un objet ou d'un groupe de objets - est moins intéressante que la fonction d'une étiquette, ou quel type d'information qu'il véhicule et comment il est utilisé. Ici, nous identifions plusieurs tags fonctions effectuent des pour les signets.

(1) L'identification de ce (ou qui) il s'agit. Une écrasante majorité, tags identifier les sujets de signet articles. Ces éléments incluent des noms communs de nombreux niveaux de spécificité, ainsi que de nombreuses bon noms, dans le cas de personnes discutant du contenu ou organisations.

(2) L'identification de ce qu'il est. Tags peut identifier ce type de chose est un élément de signet, en plus de de quoi il s'agit. Par exemple, l'article, blog et livre.

(3) Identifier à qui il appartient. Certains signets sont tagging en fonction de qui est propriétaire ou a créé le contenu en signet. Étant donné la popularité apparente des blogs parmi les utilisateurs Delicious, identifier la propriété du contenu peut être particulièrement important.

(4) catégories de raffinage. Certaines balises ne semblent pas autonome et, plutôt que d'établir des catégories eux-mêmes, d'affiner ou de qualifier des catégories existantes. Nombres, notamment autour (par exemple, 25, 100), peut exercer cette fonction.

(5) Identifier les qualités ou caractéristiques. Adjectifs comme effrayant, drôle, stupide, une source d'inspiration tag signets selon l'avis de la tagger de contenu.

(6) de référence auto. Tags en commençant par «ma», comme my stuff et my comments, identifier le contenu en termes de sa relation avec le marqueur.

(7) l'organisation de travail. Lorsque la collecte d'informations liées à l'exécution d'une tâche, que l'information pourrait être étiqueté en fonction de cette tâche, afin de grouper de ces informations ensemble. inclure le toread, jobsearch. Le groupement de l'information tâche-connexes peut être une part importante d'organisation tout en effectuant une tâche [32].

La tension entre les tags qui peuvent être utiles à la Delicious communauté au sens large et ceux qui ne sont utiles à soi-même est ici évident. Les trois premiers ne sont pas nécessairement explicitement personnels. Bien que l'identification de ce que certains points sont ou n'est présente sur certains des problèmes évoqués plus haut, comme les différences de niveau de base, ce qui unifie les quatre premières fonctions est que l'information extrinsèque à la tagger, alors on peut s'attendre à chevauchement important entre les individus. En revanche, l'élément unificateur de ces trois fonctions est définitive que l'information qu'ils fournissent est relatif à, ou seulement pertinentes pour tagger. Comme d'autres l'ont observé [46], certaines tags sont utilisées par beaucoup de gens, tandis que les autres tags sont utilisées par moins personnes. Pour les raisons décrites ci-dessus, les tags sont généralement significatif sera probablement utilisé par de nombreux tagueurs, tandis que les tags avec ou personnels spécialisés sera vraisemblablement utilisé par peu d'utilisateurs.

Les utilisateurs ont une forte tendance vers l'utilisation des tags générales première fois. Dans chaque signet, la première tag a utilisé la plus haut rang médian (soit plus grande fréquence), et successives Mots-clés ont généralement une baisse médiane de rang, à savoir la fréquence diminue (figure 3.4). Plus tôt dans la discussion des niveaux de base, une étude a montré que les principaux niveaux ont été ceux qui ont été les plus rapidement identifiés et plus généralement convenu. Nous suggérons donc que les tags plus tôt dans un signet représentent des niveaux de base, parce qu'ils ne sont pas seulement répandu dans l'accord, mais sont également les premiers termes que les utilisateurs de la pensée quand marquage des URL en question.

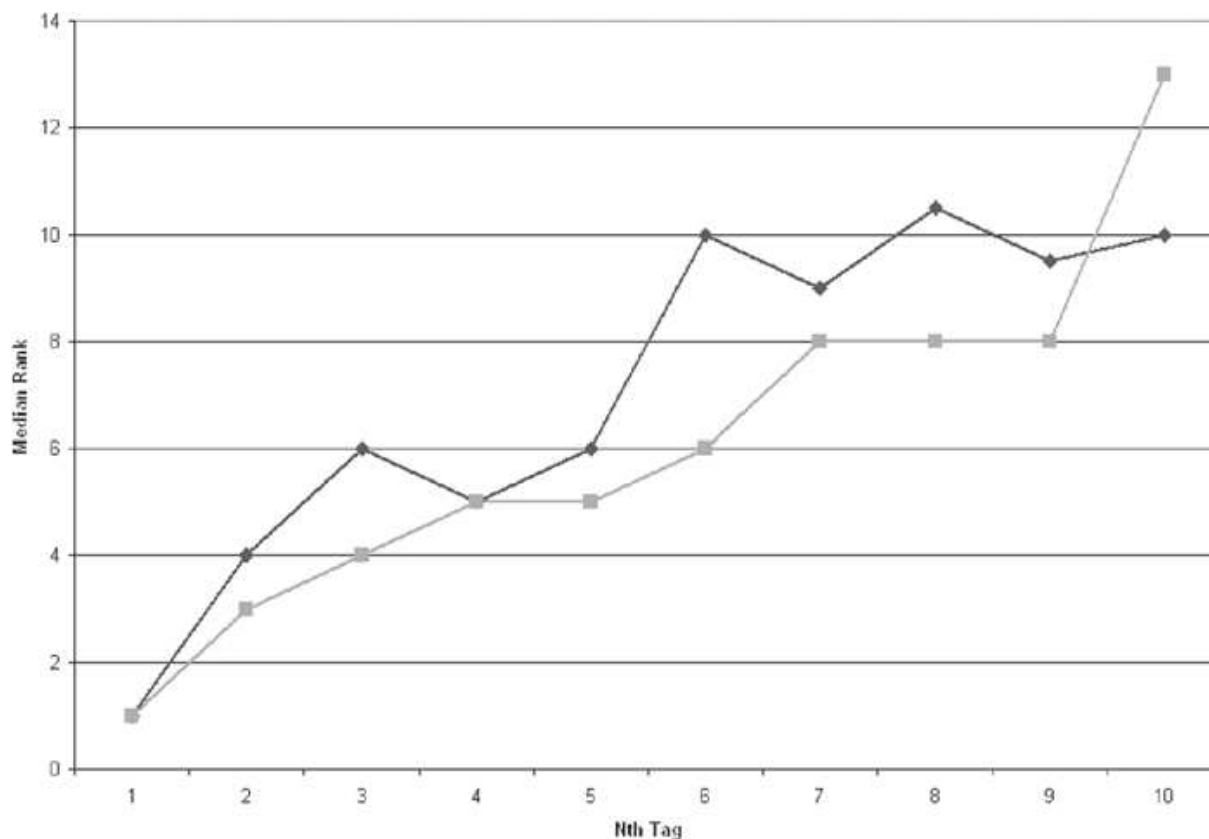


Figure 3.4 : Comme pour une balise dans un signet (horizontal) augmente, son rang (fréquence) dans la liste des tags (vertical) diminue. Cette modèle est présenté ici pour deux URL (#1209 et #1310) .

Ensuite nous allons vu quelques points forts et fiables de delicious.

3.1.3. Ses points forts :

- On ne perd plus ses signets personnels lors d'un changement de logiciel ou lors d'un remplacement d'ordinateur.
- Il simplifie la recherche des signets en évitant les menus déroulants.
- Il permet le partage de ses favoris avec d'autres utilisateurs : des amis, un collaborateur, un groupe de travail, etc.
- Il assure l'accès aux signets grâce à n'importe quel ordinateur branché à Internet.
- Il sert à bénéficier du travail d'autres internautes pour être automatiquement redirigé vers les meilleures ressources Internet sur un sujet.

3.1.4. Ses points faibles :

- Une interface assez peu ergonomique et non personnalisable :
 - ✓ Peu d'évolution depuis un an
 - ✓ Pas de possibilité de créer des dossiers
 - ✓ Affichage par date et non par thèmes
 - ✓ Pas d'insertion d'images possible
 - ✓ Graphisme peu attractif
- Ne propose pas le partage des signets par groupe d'utilisateurs
- Son principal concurrent : iGoogle/magniola/blogmark

Maintenant on va essayer de manipuler et naviguer sur delicious (comment travail dans delicious ?).

3.1.5. Utiliser Delicious sans s'y inscrire :

Delicious propose un moteur de recherche très performant sur l'ensemble des liens publiques bookmarkés par ses utilisateurs. En ce sens, delicious est un peu en concurrence avec Google qui, s'il scanne une grande partie des pages web (ce qui n'est pas le cas de delicious), recourt entre autre à des algorithmes et des calculs de machine pour donner un indice de pertinence à une page web, en fonction d'une recherche.

Tout d'abord, sans rentrer de mot-clef (appelé tag par Delicious), trois onglets proposent de savoir quels sont les derniers bookmarks (marque-pages) enregistrés, quels sont les plus populaires auprès des utilisateurs du service et enfin une exploration des bookmarks par tags.



Figure 3.5 : moteur de recherche de delicious

En tapant tag dans la barre de recherche s'affiche en dessous de cette barre tous les tags les plus utilisés similaires à votre requête. Cela est très pratique pour deux raisons. La première concerne l'orthographe d'un mot, cela permet de ne pas se tromper notamment lorsqu'il s'agit de noms propres. La deuxième permet de se rendre compte des termes utilisés par la communauté des utilisateurs de Delicious.

En effet, le modèle de tagging (manière d'écrire un tag ou mot-clef) existant sur Delicious est la folksonomie. La folksonomie est le fait de classer un élément de manière collaborative à l'aide d'un mot-clef sans que ce dernier ne soit défini par l'ensemble des utilisateurs. Ainsi, à chaque fois qu'un utilisateur du service Delicious bookmarque un nouveau lien, il peut créer un nouveaux tags et ainsi élargir le champ de recherche sur le service.

Prenons l'exemple d'une recherche sur le tag "microcredit". Delicious propose au total à ce jour **3047** résultats. Ces derniers peuvent être classés par date d'enregistrement, depuis 1 jour (bookmark très récent) jusqu'à la création du service. Delicious propose aussi une fonction liée à la folksonomie très pratique qui consiste à trier les résultats par des tags connexes (mots-clef les plus souvent associés).

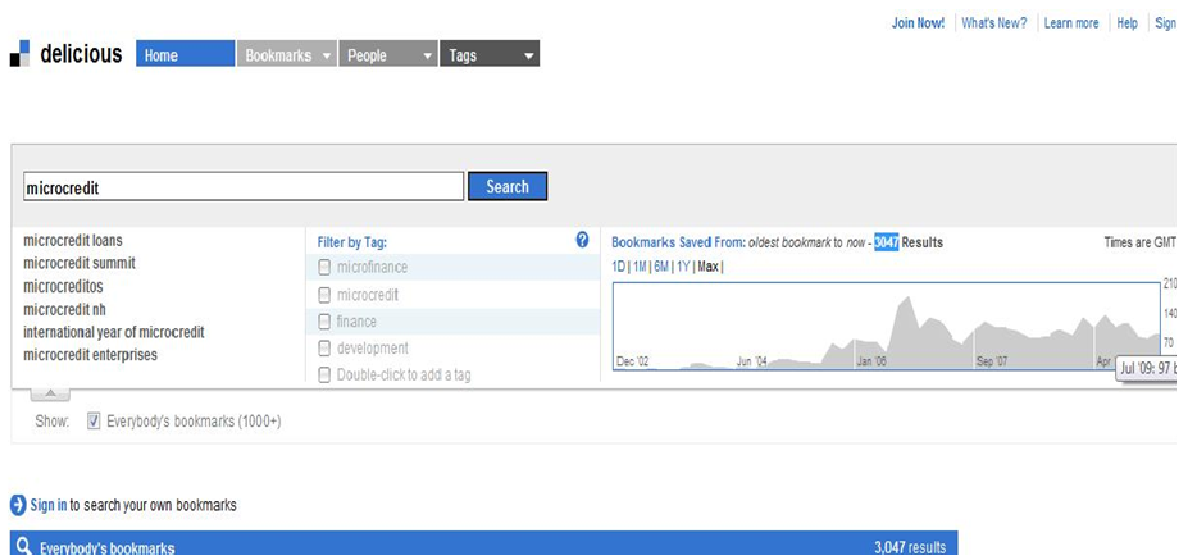


Figure 3.6 : le résultat de la recherche pour le mot « microcredit »

Des résultats communautaires :

Analysons maintenant plus précisément comment se présente le résultat d'une recherche



Figure 3.7 : un partie de résultat de recherche

Le chiffre blanc sur fond bleu en haut à gauche du résultat.

Il s'agit du nombre d'utilisateurs de Delicious qui ont aussi bookmarké ce même lien. Ainsi d'un clic sur ce chiffre apparaissent tous les utilisateurs, de celui qui a bookmarké le lien en dernier jusqu'au premier. Cela est très enrichissant, car ainsi cela ouvre de façon exponentielle le champ de recherches.

Un clic sur le pseudonyme d'un utilisateur de Delicious et il est possible de connaître l'ensemble des bookmarks (et donc des centres d'intérêts) de ce dernier. Ainsi les bookmarks de chacun fonctionnent comme des recommandations auprès des autres utilisateurs s'intéressant aux thématiques similaires. Cette démarche de recherche permet de s'ouvrir à un ensemble de bookmarks qui n'auraient pas forcément attirer l'attention de l'internaute effectuant sa recherche.

3.1.6 Création d'un compte delicious :

Nous avons maintenant de essayer de créer un compte sur delicious , pour s'inscrire, il suffit de posséder une adresse mail sur yahoo.

Etape 1 : Se rendre à l'adresse du site : <http://del.icio.us>

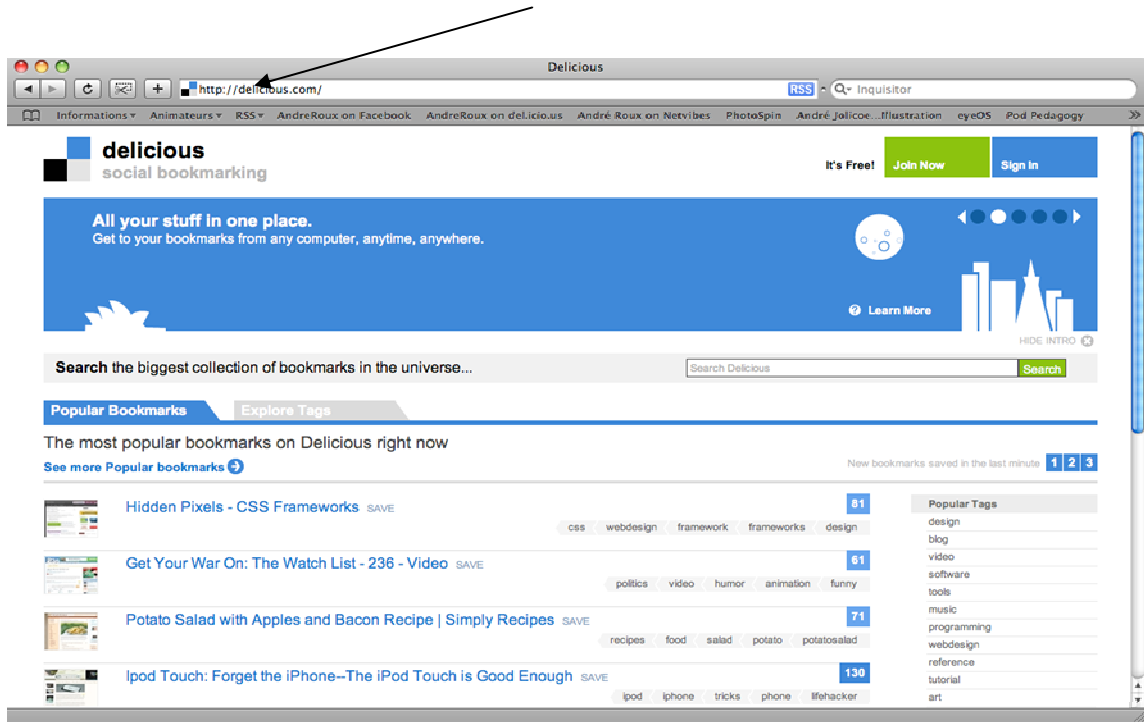


Figure 3.8 : l'interface de site delicious

Etape 2 : Pour vous inscrire, cliquez sur le bouton « Join Now », en haut à droite de la page d'accueil .

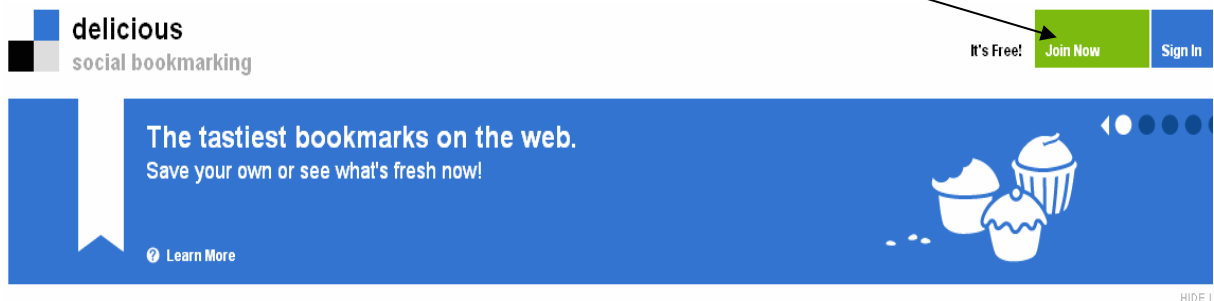


Figure 3.9 : bouton « Join Now » pour inscrire

Etape 3 : Une nouvelle page s'affiche. Remplissez les champs ci-dessous si nous avez pas un compte sur yahoo essayez de créer un compte.

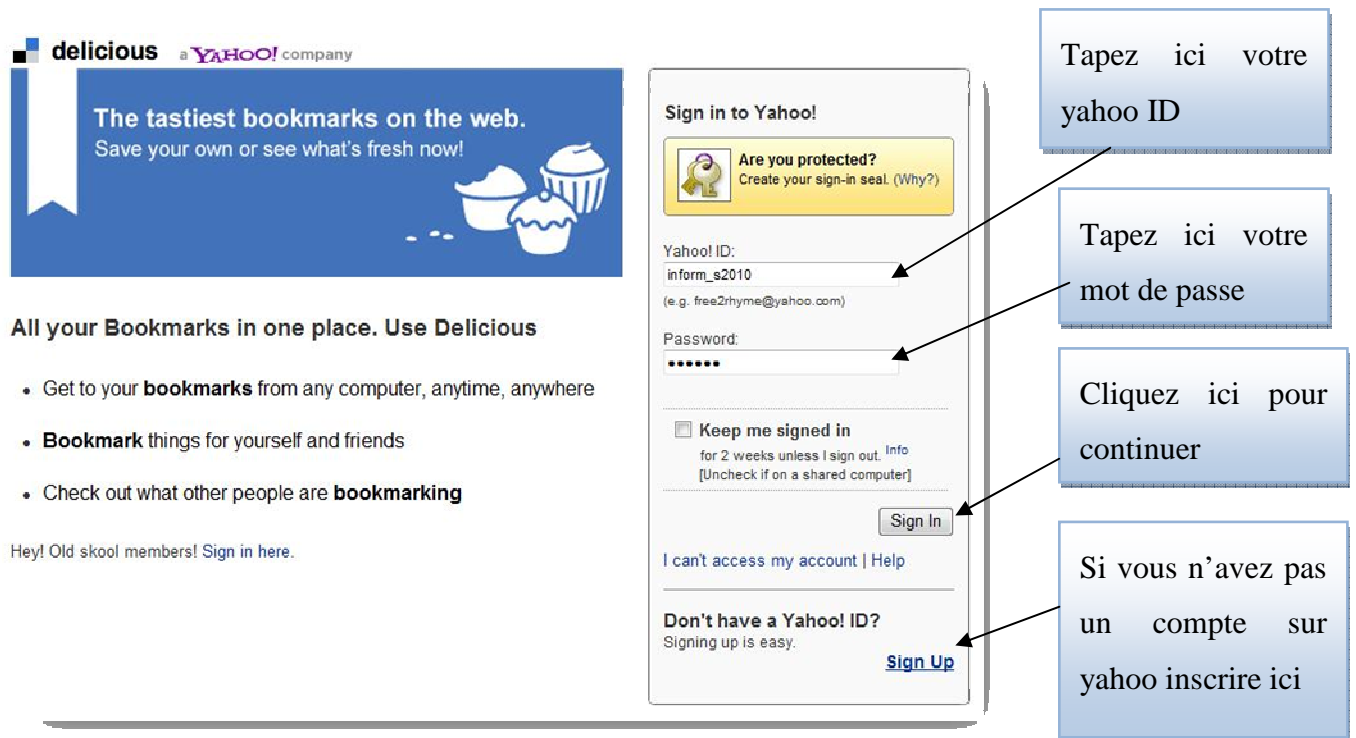


Figure 3.10 : page de logging de yahoo

Etape 4 : un nouvelle page pour continuer l’inscription

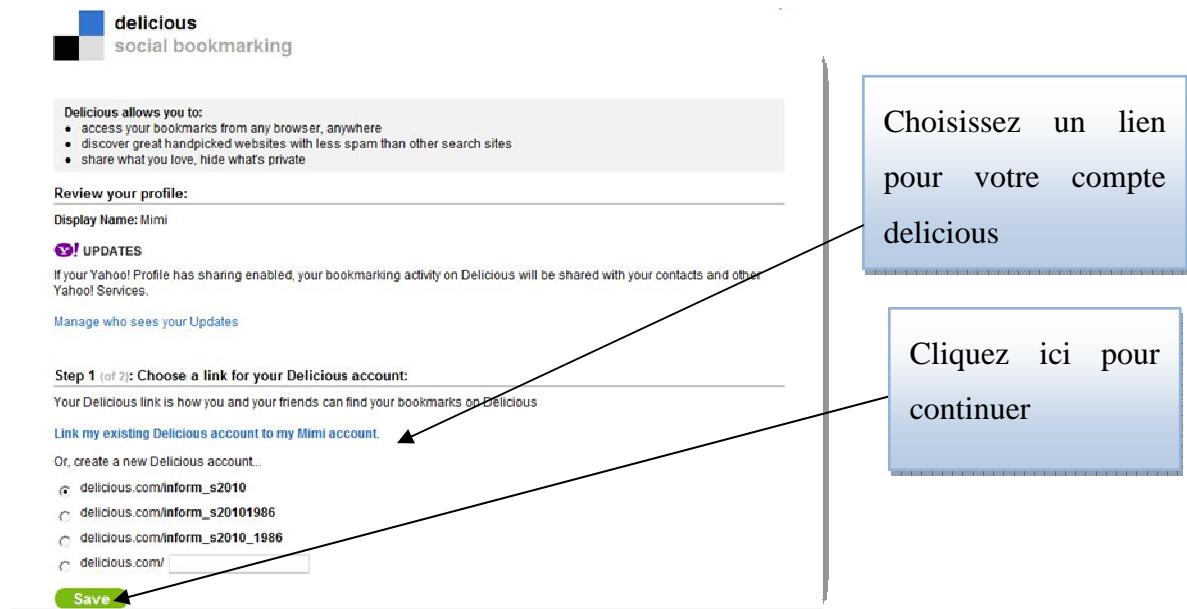


Figure 3.11 : page de confirmation de l’inscription de delicious

Etape 5 : Ajouter à mes favoris directement depuis votre navigateur avec un bookmarklet (facultatif).

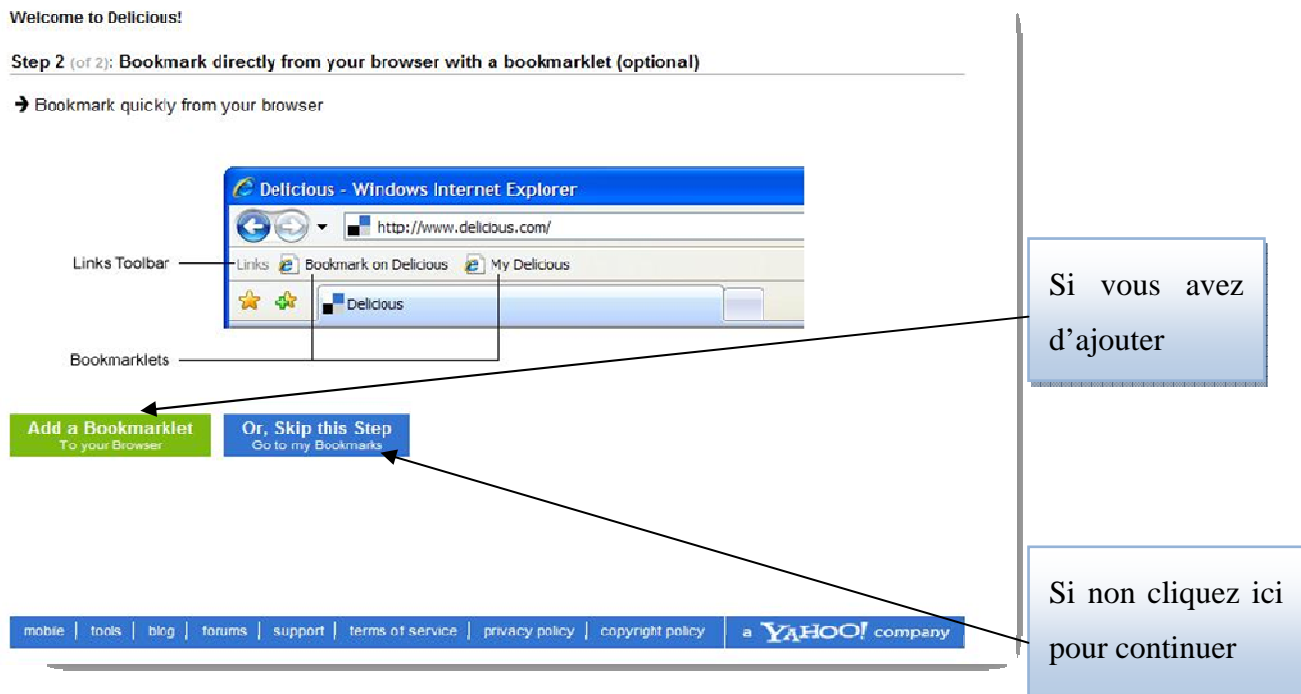


Figure 3.12 : étape d'installation des bookmarklets

Etape 6 : installer des bookmarklets

Bookmarklets (boutons) sont des liens que vous ajoutez aux favoris de votre navigateur la barre d'outils. Ils sont un moyen facile d'afficher et de consulter vos signets sur la Delicious. Seul firefox, Safari et Internet Explorer 6 et 7, Chrome et Opera sont pris en charge dans delicious

✓ **Sur Internet Explorer 7 :**

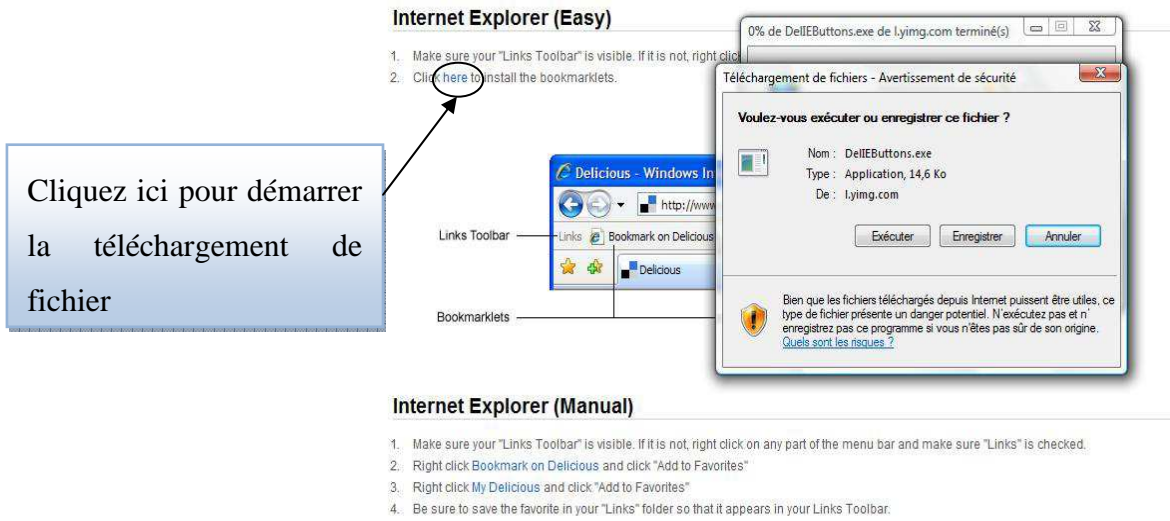


Figure 3.13 : démarrer la téléchargement de fichier

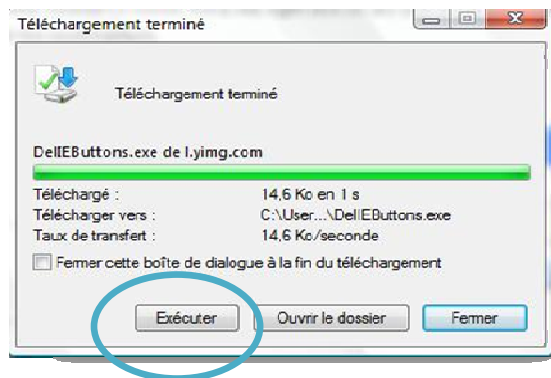


Figure 3.14 : exécuter le fichier



Figure 3.15 : démarrer l'installation



Figure 3.16 : terminer l'installation

✓ Sur Firefox :

Ajouter à mes favoris directement depuis votre navigateur avec le delicious Add-on (en option)

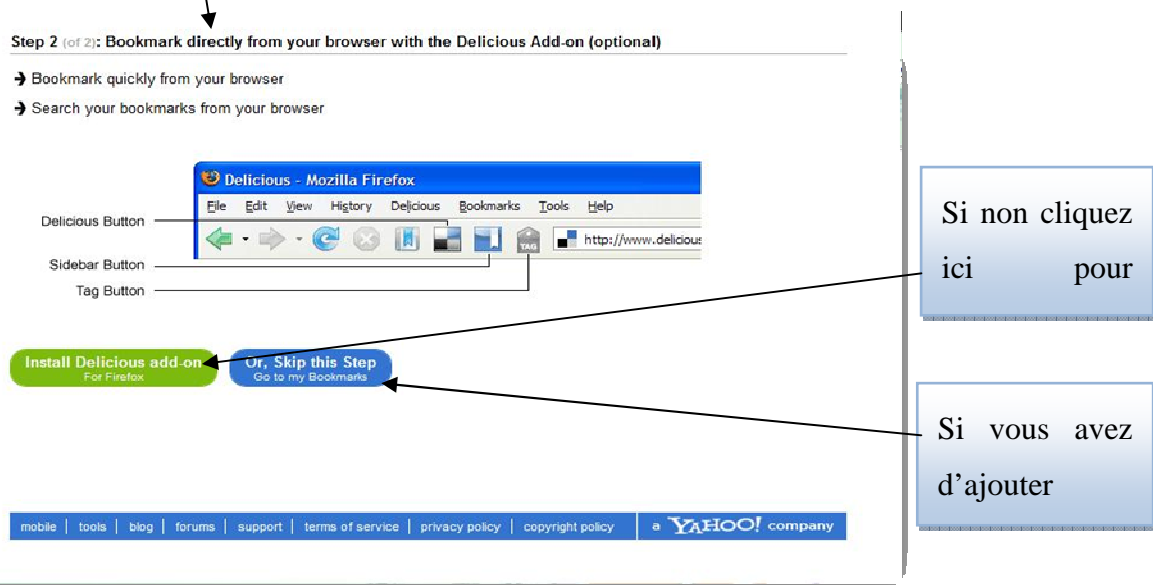


Figure 3 .17 : demande l'installation de bookmark

Firefox demande l'autorisation pour installer le logiciel



Figure 3.18 : demande l'autorisation

installer le module complémentaire

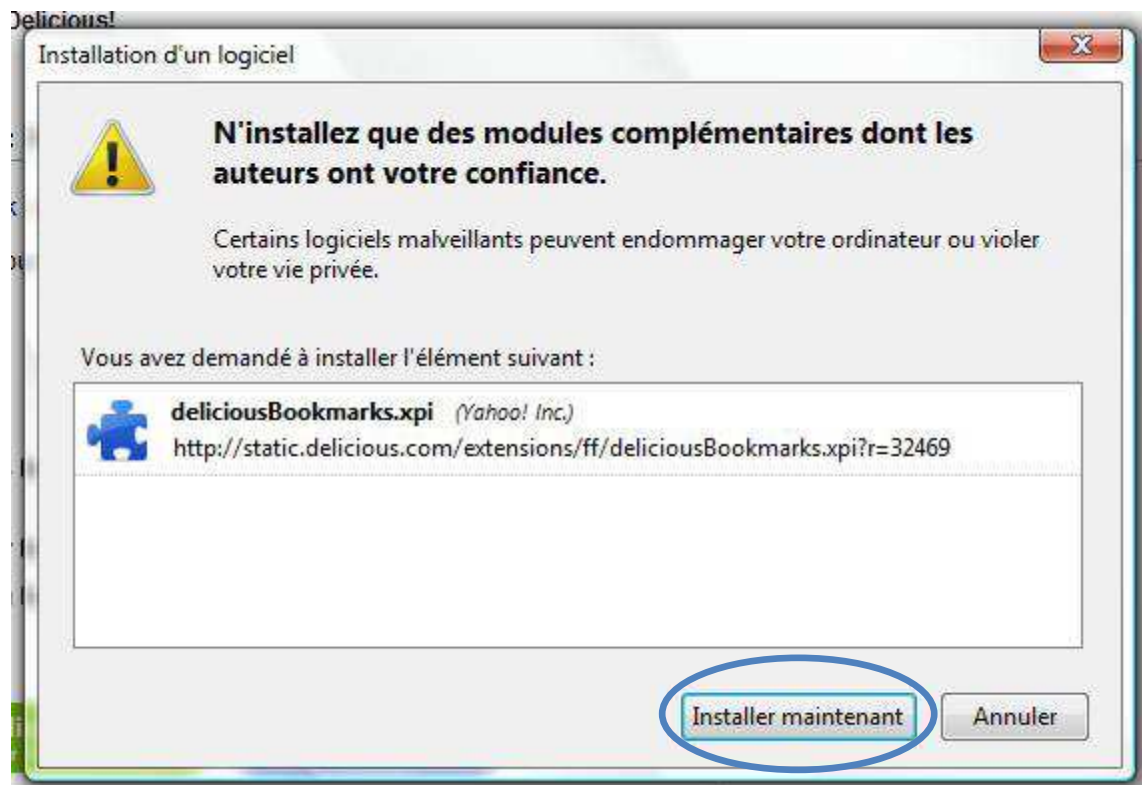


Figure 3.19 : le module complémentaire

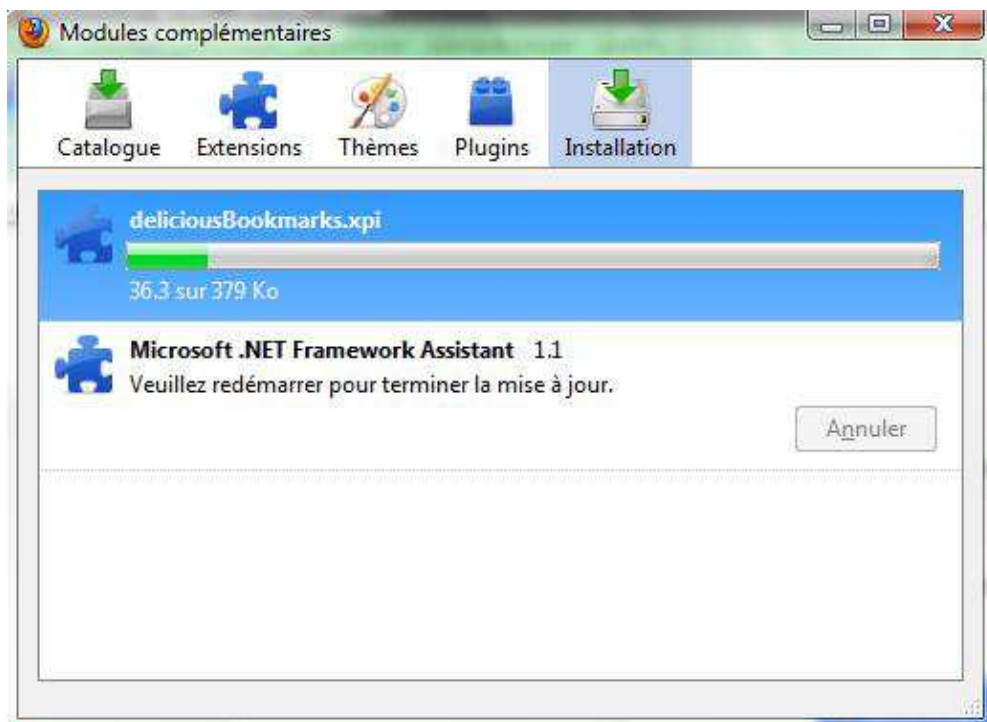


Figure 3.20 : l'installation est en cours

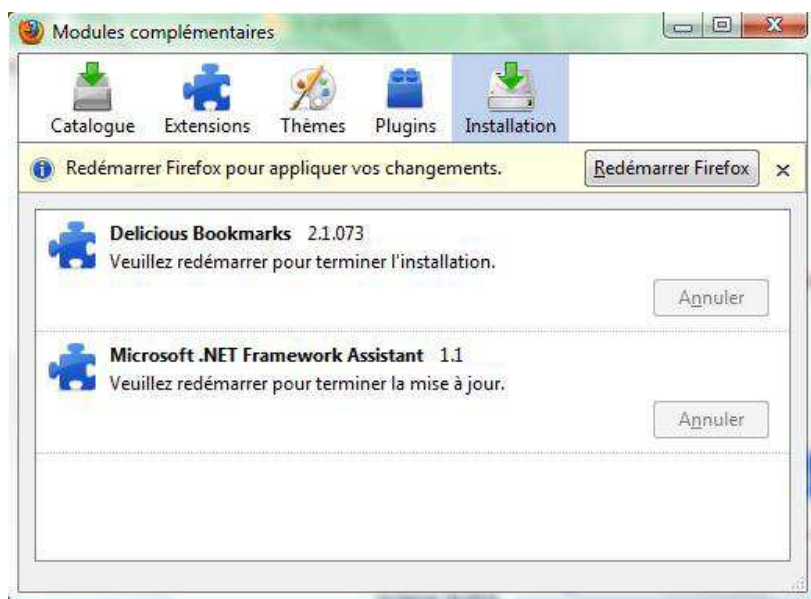


Figure 3.21: redémarrer firefox pour terminer l'installation

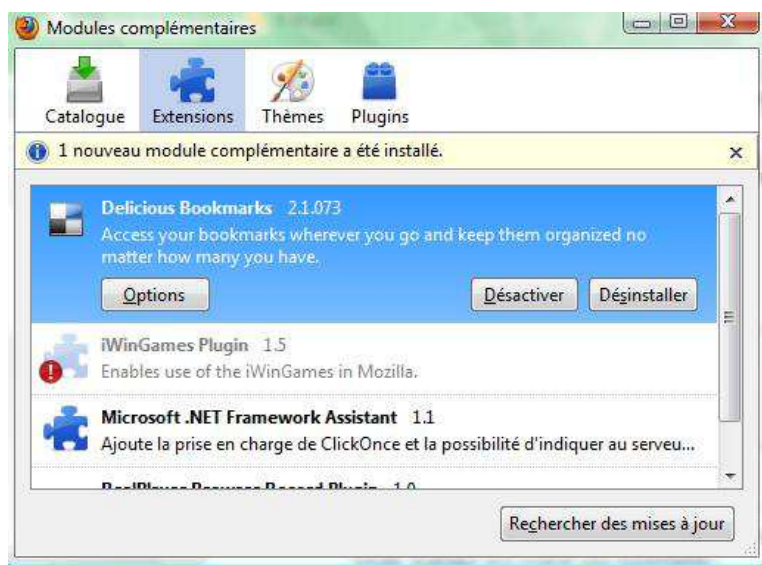


Figure 3.24 : terminé l'installation

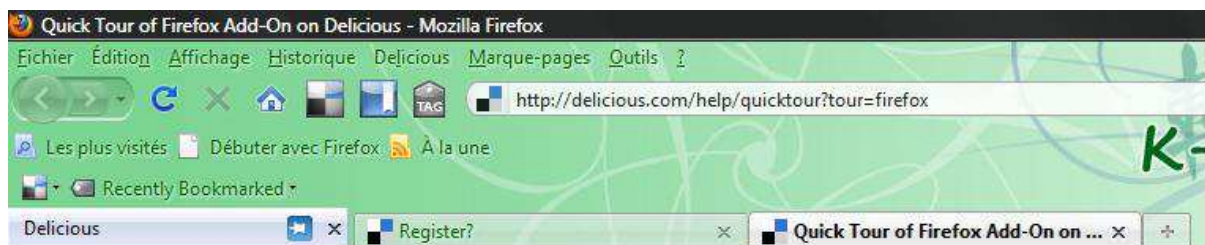


Figure 3.25 : les outils de delicious

Votre compte Del.icio.us est créé !!

3.1.7 Ajout des favoris :

Supposons dès maintenant que vous possédez un compte del.icio.us. Lorsque vous naviguez sur le web, vous êtes susceptible de vouloir mettre en favori un site que vous trouvez pertinent.

Auparavant, vous l'auriez intégré dans vos signets, maintenant vous pourrez l'inclure dans votre page Del.icio.us.

Il existe plusieurs manières pour sauvegarder, soit manuellement ou utiliser les boutons de barre de navigateur ou Importer vos favoris.

3.1.7.1 Manuellement :

sur votre compte de delicious cliquez sur « save a new bookmark »

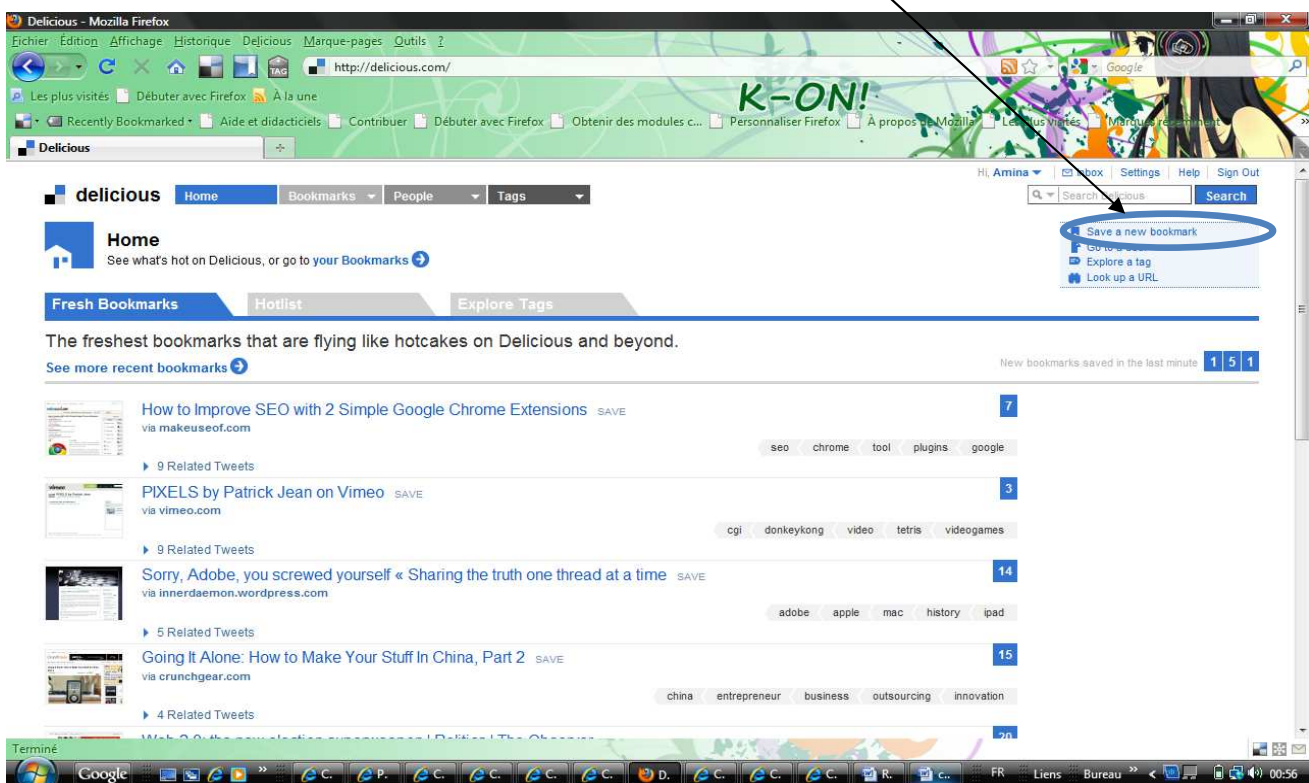


Figure 3.29 : page d'accueil

Vous pouvez copier l'adresse de la page qui vous intéresse par exemple « <http://citeseer.ist.psu.edu/> », puis la coller dans le champ « URL ».

Validez en cliquant sur le bouton « Next ».

delicious Home Bookmarks People Tags

Hi, Amina | Inbox | Settings | Search Delicious

Save a new bookmark
Start by entering a URL

Did you know? Saving bookmarks to Delicious is much easier with our [bookmarking tools](#).

URL

Sur la page suivante nommez le favori (title). souvent le nom du site apparaît automatiquement.

Appelez des tags pour retrouver plus facilement le site. Attention, un tag ne doit comporter aucun espace (si nécessaire, utilisez le tiret bas) et il doit y avoir un espace entre deux tags. Si l'annuaire est réservé au groupe, cochez « Mark as private ». Validez en cliquant sur « save ».

Save a new bookmark
Now add tags and notes

URL Required

Required

TITLE Required

NOTES 955 characters left

TAGS Space separated, 120 characters per tag

SEND

Mark as Private

Maintenant Votre site à sauvegarder !!

Les sites sont présentés sous forme de liste du plus récemment ajouté au plus ancien et ce sur plusieurs pages. Le chiffre surligné en bleu correspond au nombre de personnes qui ont ajouté le site à leur del.icio.us. Au dessous de ce chiffre, figure(nt) le ou les mot(s) clé(s) associés au site.



3.1.7.2. Utiliser les boutons de bookmark:

dans le site de cliquez sur « tag » dans la barre d'outils

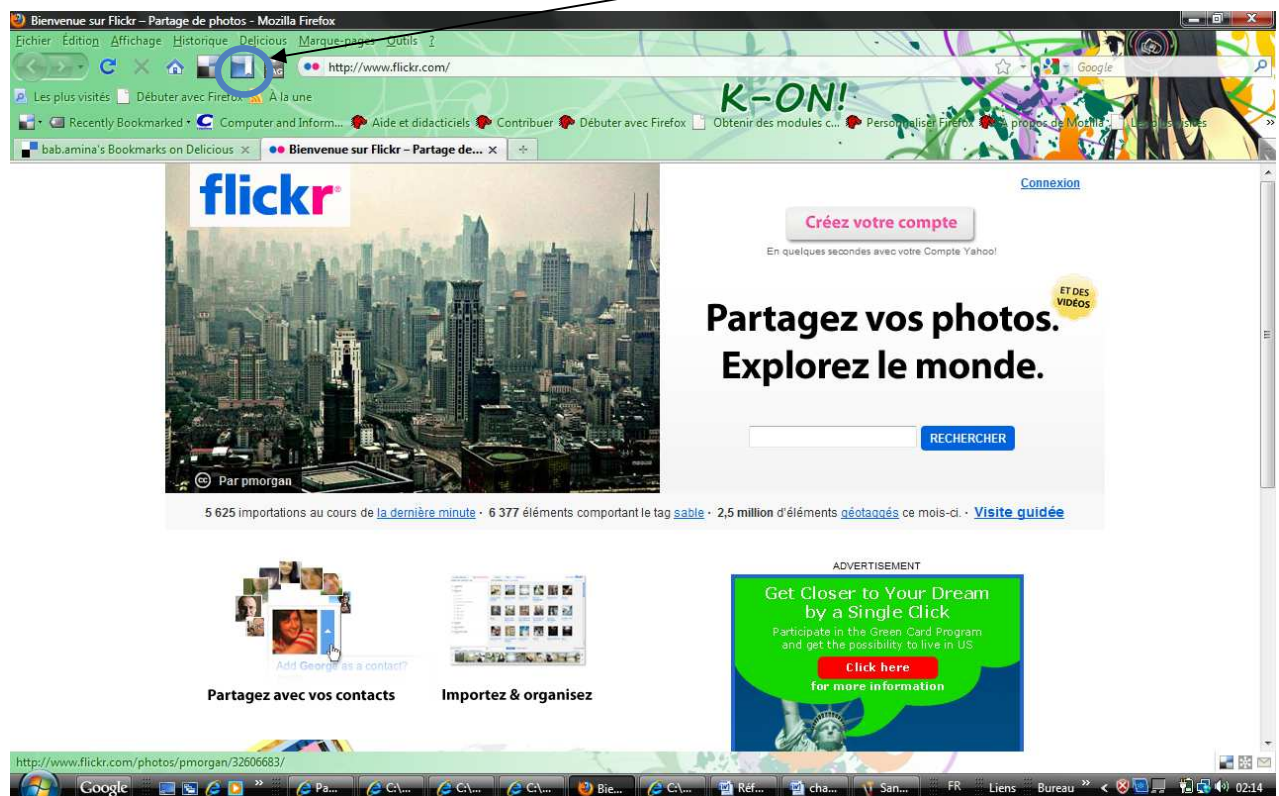


Figure 3.30: page flickr

Dans la page suivant vous remplirez le champs « tags » ou choisissez la une tag de plupart des tags proposer pour d'autre utilisateur de ce site. Si l'annuaire est réservé au groupe, cochez « Mark as private ». Validez en cliquant sur « save ».

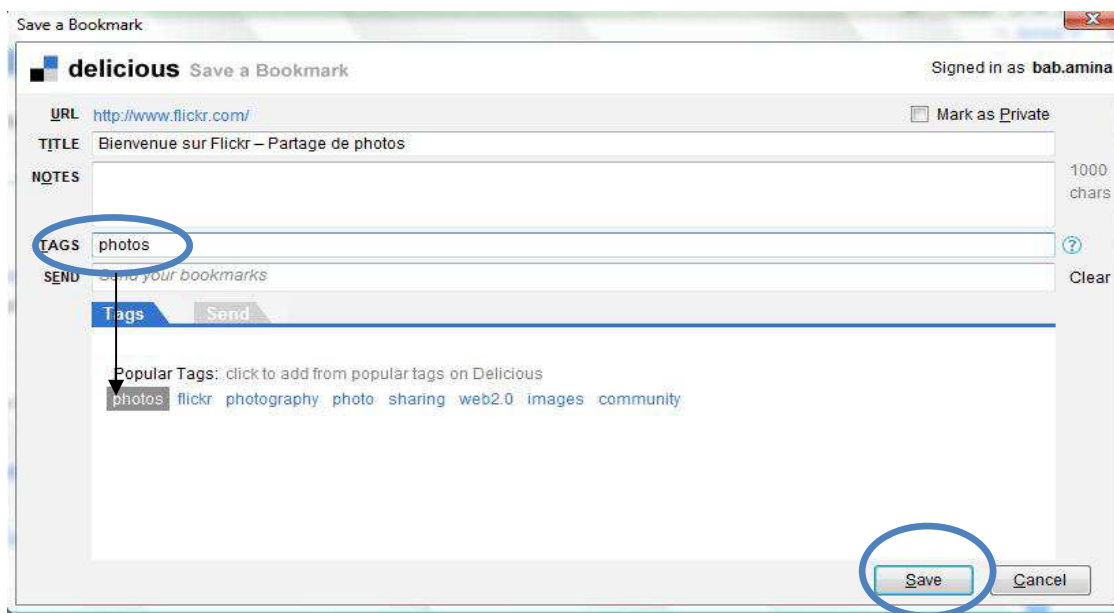


Figure 3.31 : page save a bookmark

Votre site a été favorisé !!

1.8.3. Importer vos favoris:

Pour importer votre favoris dans son compte de delicious cliquez sur « Import bookmarks » dans un votre compte.

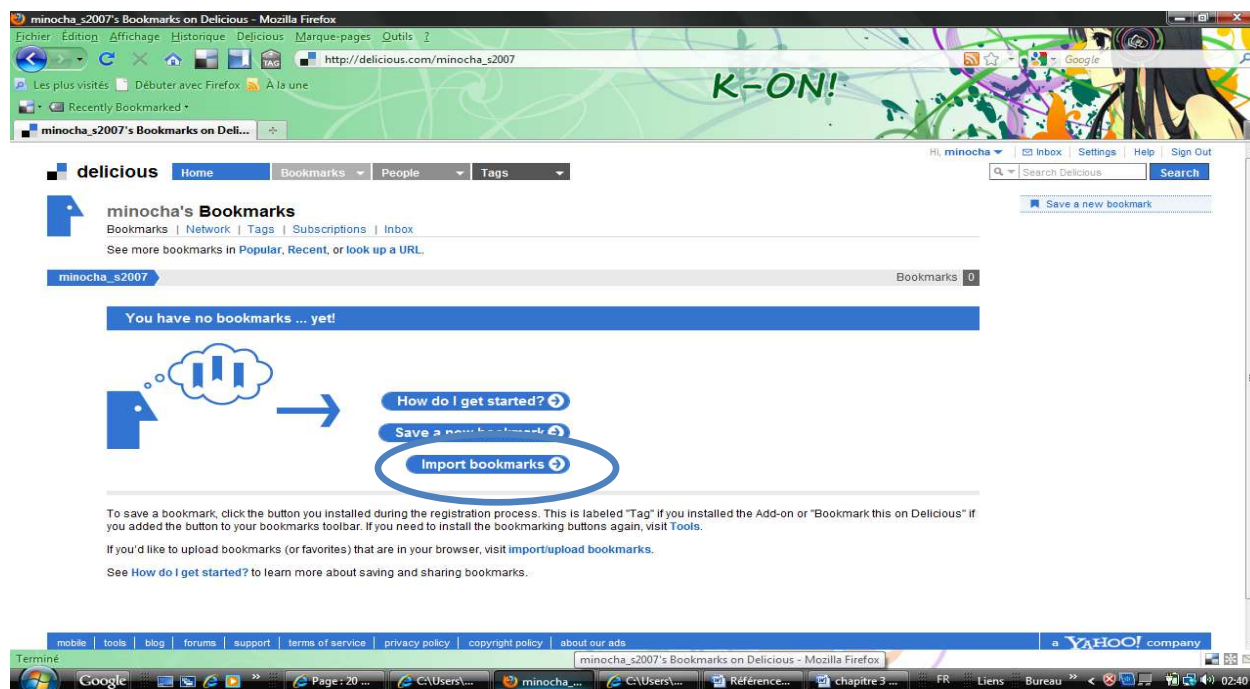


Figure 3.32 : page import bookmark

Dans la nouvelle page il existe de façon de importer votre favoris :soit « Easy Import » automatiquement importer tous ses favoris, soit « custom import » vous choisissiez votre favoris que vous importez .enfin cliquez pour « Import Now » pour terminer.

Import / Upload Your Old Bookmarks

- Easy Import** Automatically import your old bookmarks (recommended).
- Custom import** Choose your import settings...
- Email me when my import has completed.



If you have an old bookmark file you need to use our [Manual Import](#)



L’import de favoris et successif.

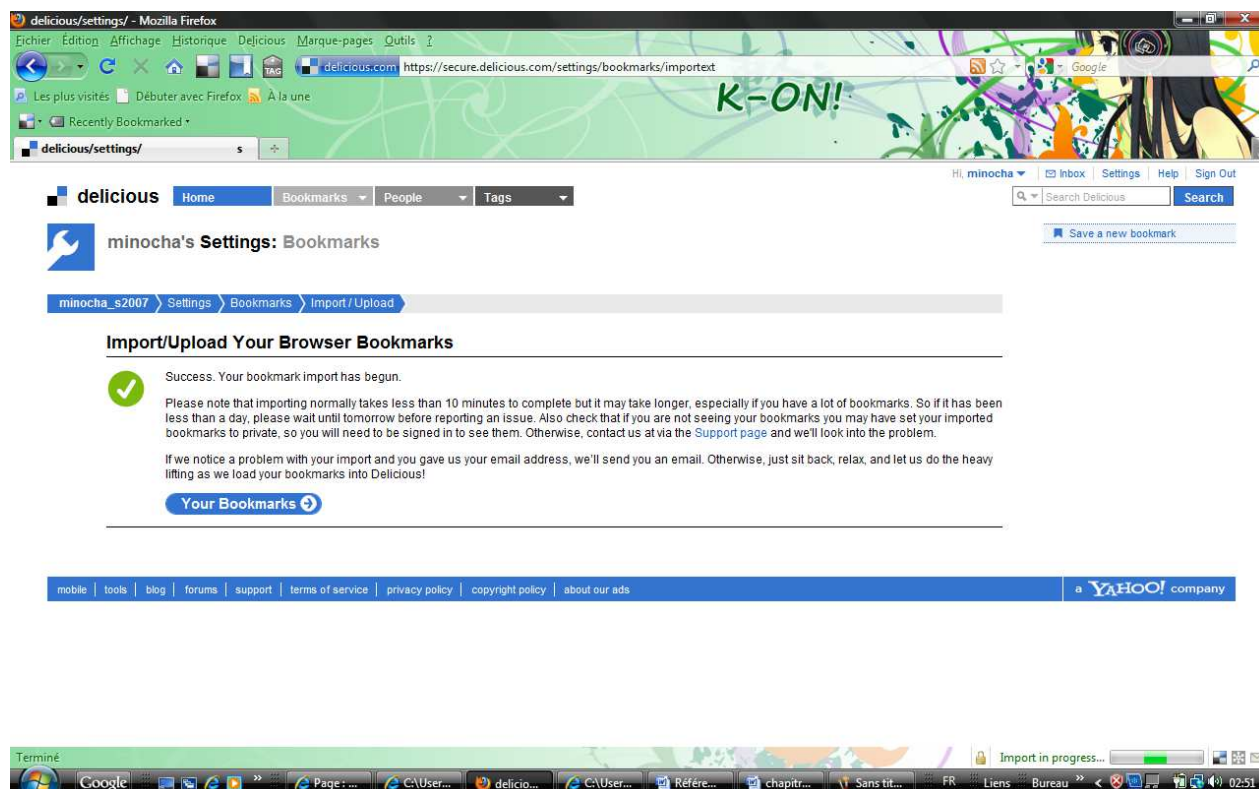


Figure 3.33 : L’importe est successif

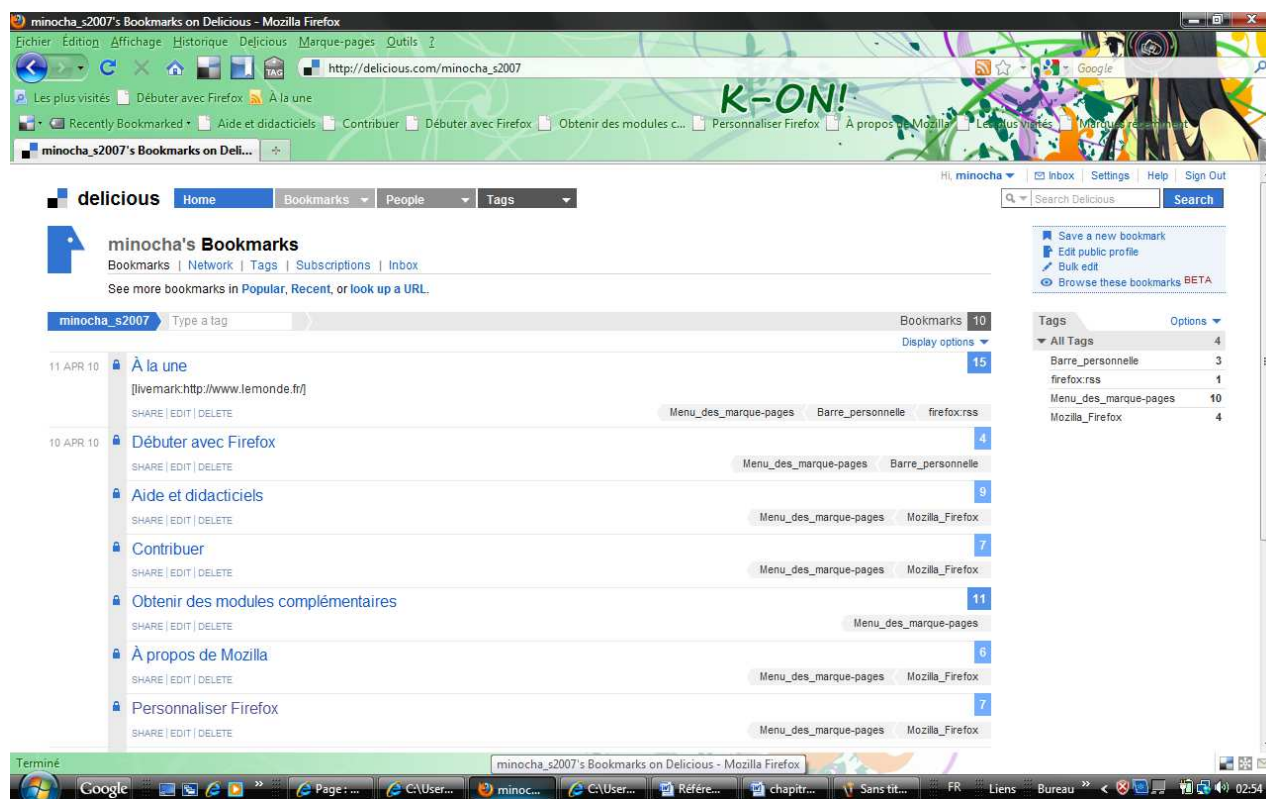


Figure 3.34 : Exemple de bookmark importé

Nous allons maintenant de essayer parler sur les tags dans delicious.

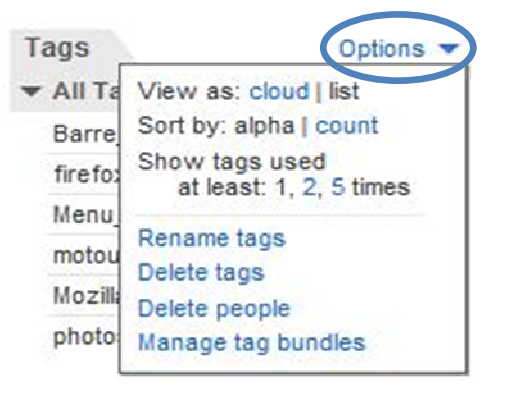
3.1.8 Organisation des tags :

Lorsque vous aurez suffisamment de favoris et que les tags se cumuleront, l’organisation des tags vous paraîtra obligatoire. Delicious permet, pour vous y retrouver, de créer des catégories de tags. Cette option vous sera indispensable pour classer vos informations. Vous pourrez croiser catégories et tags.

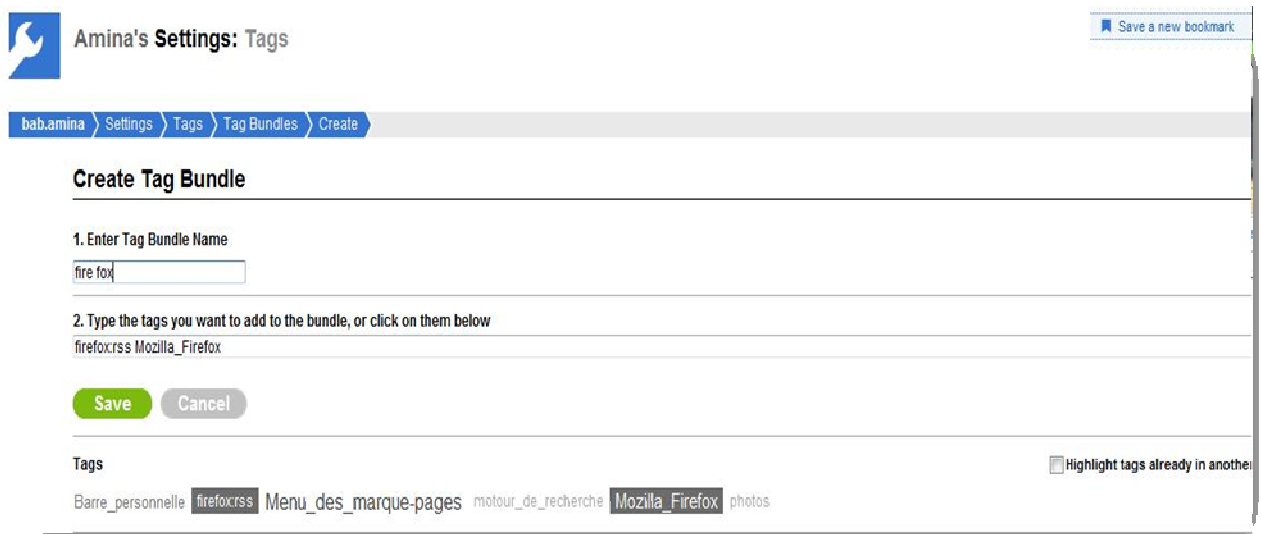
Tout se passe dans la fenêtre de tags. Cliquez sur option puis Manage Tag Bundles

Bundles Tag sont un moyen d'organiser des balises existantes en groupes. Par exemple, si vous avez les balises "design", "peinture", et "dessin", vous pouvez regrouper ces en un faisceau appelé "art"

Créez une catégorie (*Create a New bundle*) et associez à chacune les tags que vous souhaitez en cliquant dessus.



Exemple de regrouper de tags :on va regrouper le mot clé « firefox : rss » et le mot clé « mozilla_firefox » dans un nouveau mot clé « firefox »



Il est possible de gérer cette tag soit renommer ou supprimer .

3.1.9 Options d’affichage des tags :

Selon vos préférences, vous pouvez sélectionner plusieurs options d’affichage de la liste de tags dans le menu « tag options » (dans la partie de droite de la page):

· « view as cloud | list » :

- pour afficher les tags sous forme de « nuage » (cloud) : présentation « graphique », les tags sont présentés dans une police plus ou moins grande et une couleur plus ou moins foncée selon leur fréquence d’utilisation.

- pour afficher les tags sous forme de liste : liste des tags avec le nombre de favoris correspondants.



Figure 3.35 : Différent type de tags cloud .

3.1.10 Optimiser l'utilisation de Delicious dans un travail d'équipe en association ou ONG :

Delicious, outre la recherche la recherche sur internet, permet d'optimiser le partage des articles trouvés sur Internet au sein d'un même organisme.

Deux méthodes se présentent quant à l'utilisation de Delicious en équipe :

La première consiste à se créer un unique compte Delicious. Tous les membres d'un même équipe utiliseront donc un seul et unique identifiant. Cette manière de fonctionner imposera, au fur et à mesure des bookmarks enregistrés, une rigueur dans la manière de tagger chaque nouvel article.

En effet, le simple fait de tagger une étude par les tag "etude" ou "étude" (différence de l'accent sur le e) créé deux tags différents. De ce fait, lorsqu'un de vos collaboratuer rechercher tous les bookmarks que vous avez enregistré en équipe sous le tag "étude", il ne trouvera pas ceux sous le tag "etude". Il est donc très important lorque vous fonctionner en équipe sur un seul compte de bien être vigilant à la manière dont vous tagger chaque article.

Autre conseil, utilisez les "Tag Bundles". Il s'agit de "paniers" dans lesquels vous définirez les tags à ranger. Exemple : les tags "algérie", "sénégal" et "tanzanie" pourront être rangés dans un bundle "afrique".

Enfin, dernier conseil, n'utilisez que des mots-clefs et pas de préposition du type "le, la, du" etc.

La deuxième méthode consiste à ce que chacun des membres d'une même équipe se crée un compte Delicious. En fonctionnant ainsi, la première chose à faire est de se mettre en réseau. Chaque membre d'une même équipe doit s'ajouter en tant que contact Delicious. L'avantage d'une telle méthode est que le problème de l'orthographe utilisée pour le tagging disparaît mais l'inconvénient majeur vient du fait qu'un même article peut être bookmarké plusieurs fois chez différents utilisateurs d'une même équipe.

Une solution assez efficace existe pour être informé très rapidement des derniers bookmarks de ses contacts : suivre leur flux RSS Delicious. En bas de chaque page Delicious d'un utilisateur du service se trouve l'icône permettant de s'abonner au flux RSS des derniers bookmarks mis à jour.

Nous vous conseillons l'utilisation d'un seul compte au sein d'une même équipe de travail. Vous devrez définir à l'avance quels seront les principaux tags à utiliser et de quelle manière les orthographier. Un seul compte permettra aussi à votre association ou ONG d'avoir une plus forte présence sur Delicious et pourquoi pas réseauter auprès de futurs partenaires via ce nouveau canal.

On va à la suite définir un autre exemple dans le système tagging collaborative, c'est le site Flickr qui permet de partager des photos !!

3.2 Flickr :

Le deuxième système que nous avons choisi d'étudier est Flickr, un site de photo-partage populaire qui considère des étiquettes comme élément de noyau au partage, à la recherche, à la navigation, et à la découverte des images utilisateur-contribuées. Flickr permet à des utilisateurs de télécharger leurs photos personnelles à enregistrer en ligne, mais à la différence d'autres outils en ligne de photo, Flickr rend ces photos publiquement visualisables et facilement découvrable par défaut. Cette décision de conception, avec l'emphase sur étiqueter, a permis au site d'augmenter tout à fait rapidement au-dessus de sa durée de vie courte.

Flickr a été développé par la société canadienne Ludicorp en février 2004 sous la forme d'un ensemble d'outils prévus dans le cadre d'un jeu en ligne intitulé Game Neverending. Le jeu fut finalement abandonné, mais le projet Flickr issu de Game Neverending a résisté et est devenu la plateforme d'échange de photos que nous connaissons aujourd'hui. En mars 2005, Ludicorp a été rachetée par Yahoo! faisant de Flickr l'un de ses services.

Flickr est gratuit pour la majorité de ses fonctions mais l'accès à certaines fonctionnalités avancées est toutefois soumis au paiement d'un abonnement annuel.

Depuis avril 2008, Flickr accepte, en plus des photos, les vidéos d'une durée maximale de 90 secondes.

En novembre 2008, Flickr a annoncé avoir dépassé les 3 milliards de photos uploadées sur le site et réunir plus de 50 millions d'utilisateurs uniques.

Bien que cet outil soit très pratique et utile pour présenter ses meilleures oeuvres photographiques au monde entier, il est important de l'utiliser à bon escient et de se méfier des risques d'atteinte à la vie privée, au droit à l'image et au droit d'auteur, si Flickr est mal utilisé.

Cette croissance a en partie été dû à la grande sélection de social supports de Flickr d'interactions : en plus des photos de chargement, les utilisateurs peuvent également créer des réseaux des amis, joignent des groupes, les envoient message à d'autres utilisateurs, commentaire sur des photos, photos tags, choisissent leurs photos favorites, et ainsi de suite. Cette abondance d'outils de transmission et les formes de l'organisation sociale crée l'écologie

interconnectée de médias d'une façon forte - qui peut mener des utilisateurs aux personnes éloignées et des places avec seulement quelques clics. Les tags sont une part importante de cet environnement, où les tags agissent en tant qu'outil de navigation primaire pour trouver les ressources et les personnes semblables.

Comme précédemment remarquable, l'analyse la plus étendue d'un système de tagging a été terminée sur des données rassemblées du social site bookmarking Del.icio.us [9]. Nous avons choisi Flickr à fournir une traduction alternative aux conclusions dérivées de cette étude. Dans presque chaque catégorie dans notre taxonomie de système, Flickr occupe un espace alternatif de Del.icio.us : il contient les ressources utilisateur-contribuées par opposition à global ; tagging des droites sont limitées au l'individu-tagging (et au mieux permission-basées, bien que dans la pratique individu-tagging dans plus répandu) au lieu d'un général ; des tags sont agrégées dans les positionnements au lieu des sacs ; et en conclusion, l'interface a les moyens la plupart du temps pour aveugle-tagging au lieu du suggérer-tagging.

Ces décisions de conception forment les structures encourageantes qui pilotent des personnes aux ressources tag. Puisque Del.icio.us tâche-est en grande partie focalisé, à savoir enregistrer bookmarks pour la future recherche, des motivations d'organisation sont le plus dominant. Tandis que l'élément social de tagging est évident du leveraging de la communauté la contribution, un manque de systèmes de communication (par exemple transmission de messages ou réseaux sociaux explicites) de emphatise des incitations sociales non-d'organisation.

Les utilisateurs de Flickr, d'une part, sont susceptibles également tagging pour leur propre recherche, mais couplé avec une abondance de transmission les mécanismes, la conception de système encourage le jeu et exploration d'utilisation de tag. Des utilisateurs sont principalement motivés par des incitations sociales, y compris les occasions de partager et jouer.

Dans l'analyse suivante nous présentons une analyse d'utilisation de tag dans Flickr pour proposer des champs de recherche potentiels pour la future étude.

Nous avons eu l'occasion de fonctionner directement avec un sous-ensemble de la base de données employée par Flickr, spécifiquement informations sur des photos, des tags, et les

rapports sociaux explicites entre les utilisateurs (c.-à-d., le réseau de « contact »). Puisque notre foyer est sur l'utilisation des tags, nous avons choisi seulement ces utilisateurs qui ont utilisé ce dispositif (c.-à-d., utilisé au moins une tag pour décrire une photo) et seulement ces photos qui ont eu au moins une tag appliquée. Des millions d'utilisateurs de Flickr, nous avons aléatoirement choisi un ensemble de 25.000 pour notre analyse de différents comportements ; pour le cas plus compliqué de l'analyse réseau, nous avons choisi un autre sous-ensemble de 2.500.

Cette étude est seulement un regard préliminaire à la dynamique du système de Flickr et est censée exposer des tendances et des sujets intéressants dans les données de Flickr. Ces sujets illustrent de divers aspects d'étiqueter les systèmes et leur structure encourageante, mais nous n'essayons pas de ne prouver ou affirmer aucune conclusion générale au sujet de tous les systèmes de tagging.

Nous avons eu l'occasion de fonctionner directement avec un sous-ensemble de la base de données employée par Flickr, spécifiquement informations sur des photos, des tags, et les rapports sociaux explicites entre les utilisateurs (c.-à-d., le réseau de « contact »). Puisque notre foyer est sur l'utilisation des tags, nous avons choisi seulement ces utilisateurs qui ont utilisé ce dispositif (c.-à-d., utilisé au moins une tag pour décrire une photo) et seulement ces photos qui ont eu au moins une tag appliquée. Des millions d'utilisateurs de Flickr, nous avons aléatoirement choisi un ensemble de 25.000 pour notre analyse de différents comportements ; pour le cas plus compliqué de l'analyse réseau, nous avons choisi un autre sous-ensemble de 2.500.

Cette étude est seulement un regard préliminaire à la dynamique du système de Flickr et est censée exposer des tendances et des sujets intéressants dans les données de Flickr. Ces sujets illustrent de divers aspects tagging les systèmes et leur structure encourageante, mais nous n'essayons pas de ne prouver ou affirmer aucune conclusion générale au sujet de tous les systèmes de tagging.

3.2.1. Utilisation de tag :

Les tags ne sont pas obligatoires dans le modèle d'utilisation de Flickr. Dans un système de tagging social, les tags sont typiquement une fonctionnalité optionnelle dans une plus grande tâche d'organisation de ressource. Comme Del.icio.us, l'interface de Flickr incite des utilisateurs pour des méta-données au sujet de chaque ressource identifiée : un titre, une légende, et une liste de tags. Dans le cas des deux systèmes, la puissance d'entrée de tag vient troisièmement dans l'interface de puissance d'entrée, mais les différencie également d'autres outils de gestion des ressources.

En plus de tagging ses propres photos, le système de Flickr permet également à des utilisateurs photos de tags leurs des amis. Cependant, ce dispositif n'est pas en grande partie utilisé ; de 58 millions de tag nous avons observé, seulement un petit sous-ensemble sont de ce type ; une majorité écrasante des tags sont appliquées par les propriétaires des photos.

Les configurations d'utilisation d'étiquette varient tout à fait rigoureusement parmi des utilisateurs de Flickr, et comme prévu, fait ainsi l'adoption d'étiqueter le comportement. Le figure 40 affiche la fonction de répartition cumulative (CDF) pour le tag taille de vocabulaire à travers l'ensemble d'utilisateurs. La valeur à une valeur indiquée est la probabilité (axe des ordonnées) qu'un utilisateur aléatoire a un ensemble des tags distincts (axe des abscisses) qui est plus grand que cette taille de collection. Par exemple, la probabilité qu'un utilisateur de Flickr a plus de 750 tags distincts est approximativement 0.1%. Cette distribution illustre le fait que la plupart des utilisateurs ont très peu des tags distincts tandis qu'un petit groupe a les ensembles des tags extrêmement grands.

Le rapport entre l'utilisation de tag et d'autres types de puissance d'entrée peut être un bon indicateur de la façon dont les utilisateurs utiles ou importants croient que les tags sont à l'expérience d'utiliser le système. Dans Del.icio.us, Golder et Huberman ont constaté qu'il n'y avait pas un fort association entre le nombre de signets faits et numéro des étiquettes employées pour annoter ces signets [28]. Nous avons étudié trois activités dans l'environnement de Flickr : le nombre de photos téléchargées, le compte des tags distincts de l'utilisateur, et numéro des contacts indiqués par l'utilisateur. Par exemple, un certain l'utilisateur peut avoir 100 photos avec un total de 200 tags distincts à travers ces photos, et soient connectées à 50 contacts différents.

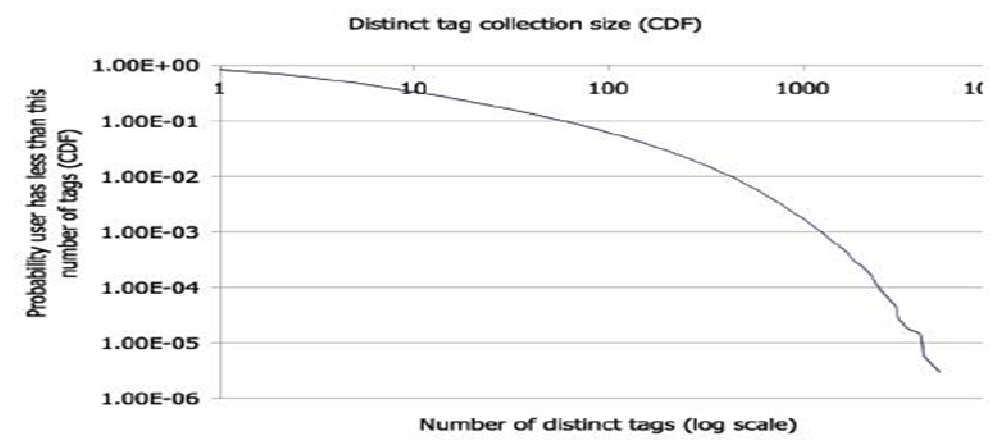


Figure 3.36 : Distribution (densité) du numéro des tags distincts par utilisateur

Le tableau 2 affiche par paires la corrélation de Pearson [8] entre taille de collection de photos, tags distincts et numéro des contacts à travers l'ensemble d'utilisateurs. Nous avons calculé cette corrélation pour un ensemble de 25.000 utilisateurs aléatoirement choisis parmi notre ensemble de données. Par exemple, la corrélation entre les tags et les photos est 0.518, suggérant un rapport linéaire fort entre ces variables, c.-à-d. une augmentation de taille de collection de photos implique une augmentation du nombre de tags distincts. Le rapport le plus fort entre ces trois éléments (photos, tags distinctes, et contacts) vient entre les photos et distinct tags, un rapport probable étant donné que le tagging de ceux' pour posséder des photos est la forme dominante des tags.

Tableau2 : Corrélation d'utilisation de Flickr

	Tags	Photos	Contactes
Tags	1	.518	.386
Photos	.518	1	.192
Contactes	.386	.192	1

* N = 25.000

** p < 0.001 pour toutes les valeurs.

L'association entre les contacts et les photos est beaucoup plus faible que celui entre les contacts et les étiquettes distinctes, qui pourraient suggérer que le tagging soit lié à l'activité sociale à un certain degré.

En plus des implications sociales, un autre dispositif des tags intéressant étudier est l'utilisation d'un individu des tags avec le temps. Comment la fréquence des tags change-t-elle pendant qu'un utilisateur devient acclimaté au système ? Faire ses tags deviennent une taxonomie cohésive avec le temps, ou continuent-elles à se développer pendant que ses expériences changent ? En étudiant Del.icio.us, Golder et Huberman affichent des exemples quant à la façon dont ensembles de certains utilisateurs les' d'étiquettes distinctes continuent à se développer linéairement pendant que de nouvelles ressources sont ajoutées. En même temps, ils réclament que les étiquettes pour une ressource donnée tendent à stabiliser après que seulement quelques utilisateurs aient étiqueté lui [9]. Puisque Flickr utilise un placer-modèle pour représenter des tags, nous ne pouvons pas réexaminer la dernière observation, mais nous pouvons regarder la croissance des tags d'un utilisateur avec le temps.

Nous allons maintenant faire un peu de pratiquer sur flickr .

3.2.2 Créer un compte sur flickr :

Flickr étant un service de Yahoo!, il est requis de créer un compte e-mail Yahoo! avant d'ouvrir un compte Flickr. Le mot de passe pour accéder à Flickr devient celui de sa messagerie Yahoo.

Si vous avez déjà un compte de messagerie Yahoo ! Ou Ymail, vous pouvez utiliser Flickr sans avoir à créer un nouveau compte, il vous suffit alors de vous rendre directement dans la Connexion au compte Flickr. Si vous ne possédez pas encore de compte Yahoo, vous devrez en créer un compte Rendez vous sur le site de Flickr (<http://www.flickr.com/>) pour créer un compte.

Etape 1 : Cliquez sur le bouton « **Créez votre compte** ». Suivez la procédure en remplissant le formulaire.

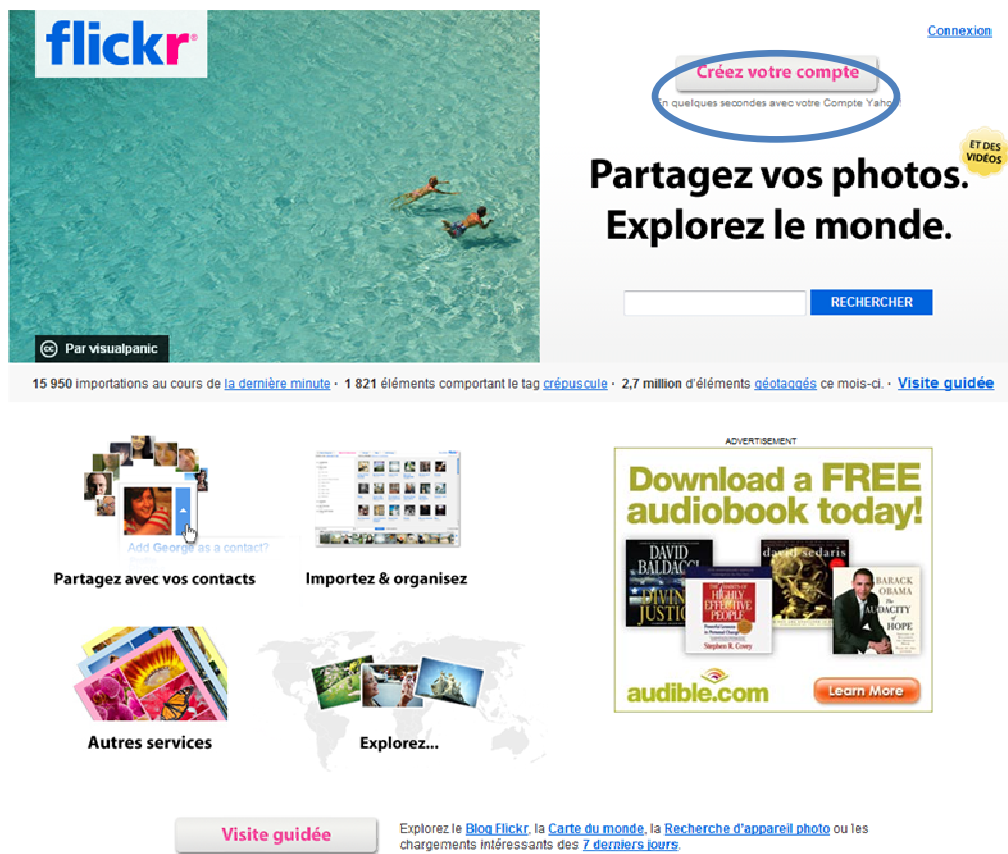


Figure 3.37 : Page d'accueil de Flickr

Etape 2 : saisir votre adresse email de yahoo et son mot de passe puis cliquer sur le bouton « ouvrir une session » si vous n'avez pas un compte sur yahoo cliquez sur « je m'inscrire »

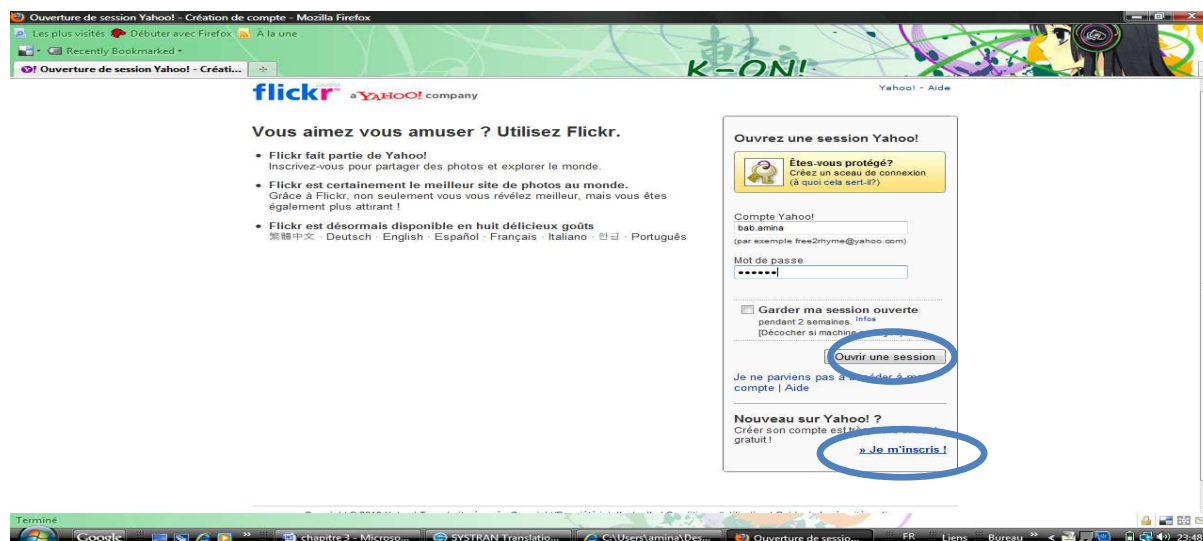


Figure 3.38 : page ouvrir une session

Etape 3 : S’il s’agit de la première connexion sur son compte Flickr, elle sera redirigée vers Flickr où elle doit choisir un pseudo. Il peut être différent de celui du compte Yahoo!

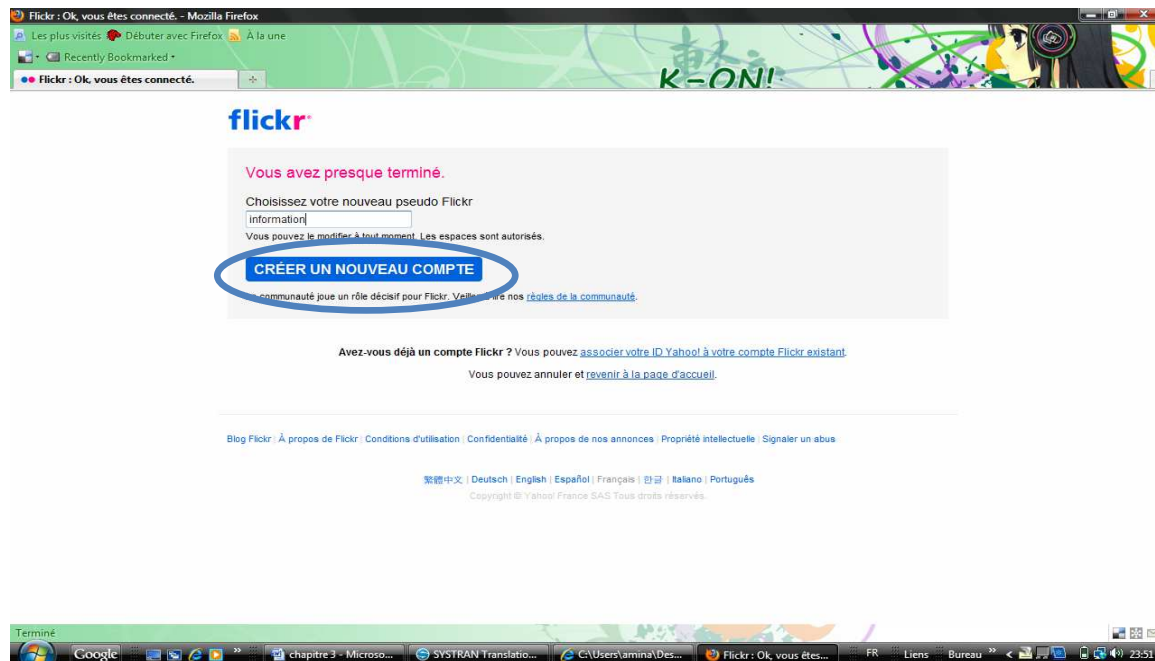


Figure 3.39 : Personnalisation du compte Flickr

Une fois le pseudonyme choisi, pourra profiter à son compte Flickr.

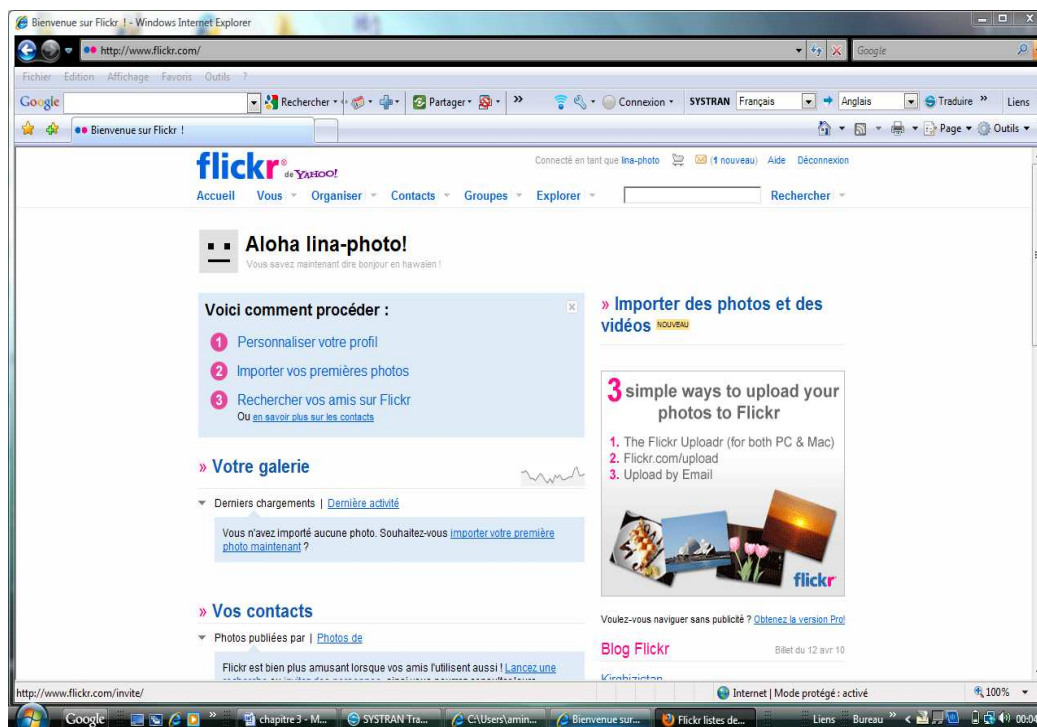


Figure 3.40 : Page d’accueil personnalisée de lina-photo

Maintenant flickr permet de importer vous photo !!

3.2.3 Importer les photos dans ce compte de flickr :

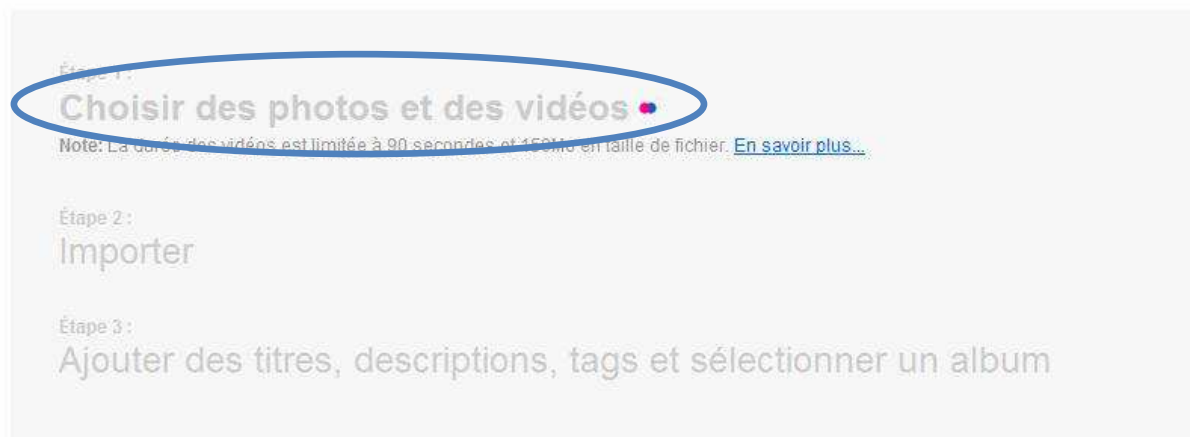
Pour envoyer des photos sur le serveur Flickr cliquez sur le lien **Importez vos premières photos** sur la page d'accueil de votre compte Flickr.



Puis cliquez sur « **Choisir des photos et des vidéos** »

Importer vers Flickr

Vous avez utilisé 0% de votre limite en photo de 100,00 MB et 0 de votre limite en vidéo de 2 pour ce mois-ci. [Mettre à niveau ?](#)



Vous pouvez télécharger plusieurs images à la fois. Pour cela, cliquez pour chacune d'entre elles sur le bouton **Parcourir** et choisissez sur votre ordinateur les images souhaitées.

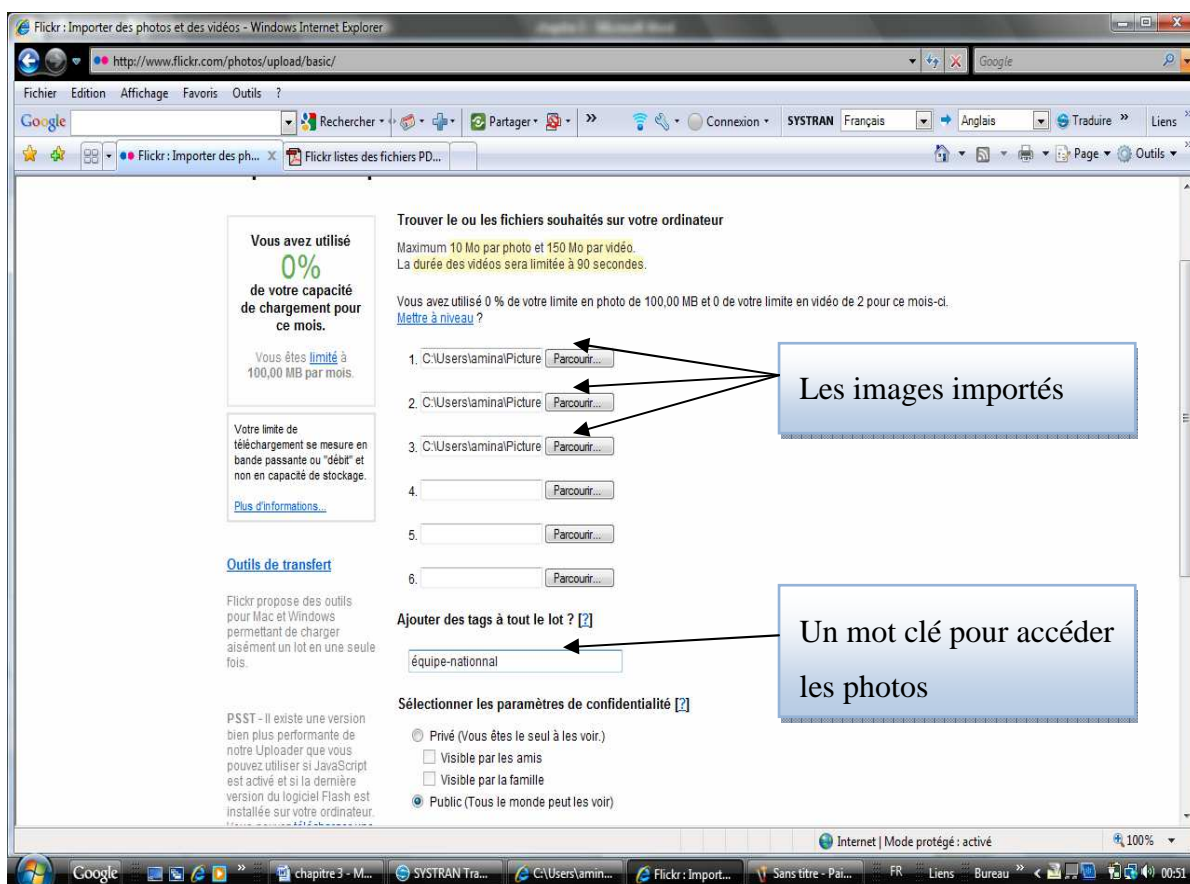


Figure 3.41 : Page de parcourir le photo

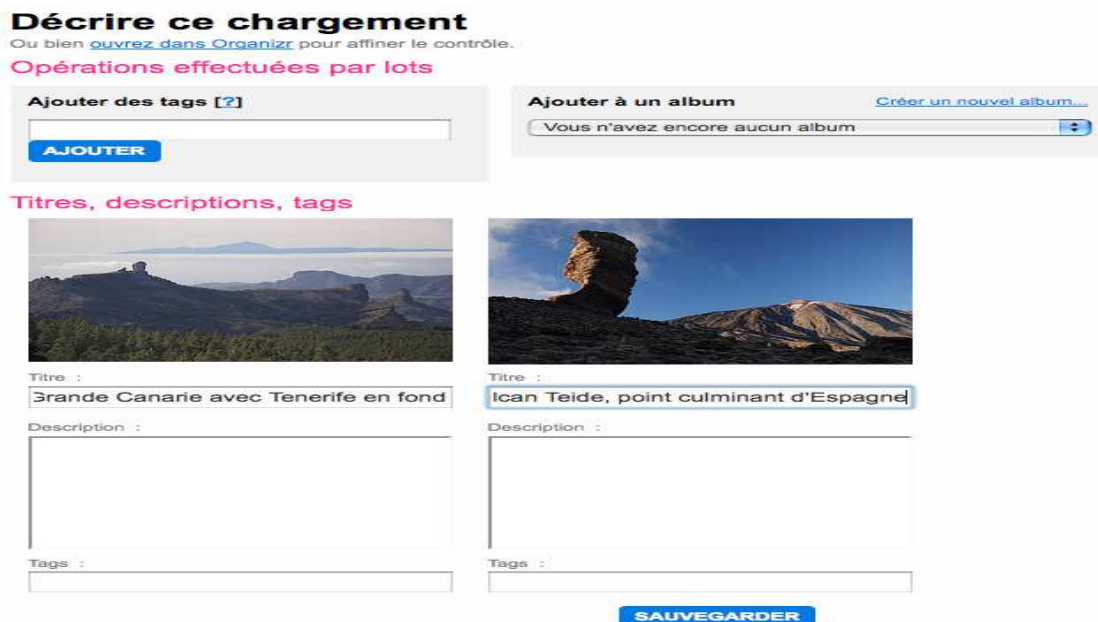


Figure 3.42 : Les photos importer

Vous pouvez également classer vos photos en album de façon très simple.

Autres moyens de charger vos photos : vous pouvez également télécharger vos photos sur le serveur par le biais d'une adresse e-mail qui vous sera attribuée ou utiliser un a application que vous installerez sur votre ordinateur.

3.3. YouTube :



Aujourd'hui, les vidéos sont incontournables sur le Web. Les grands journaux en agrémentent leurs pages, les chaînes spécialisées se développent pour partager du contenu audio et vidéo, professionnel comme amateur, et les internautes ne sont pas en reste car ils sont les plus gros consommateurs et producteurs de vidéos. Youtube, principalement consulté par un public anglophone, en est encore le leader mondial.

Beaucoup de services professionnels se dégagent et produisent des émissions indépendantes, du contenu sponsorisés avec ou sans publicité et offrent des plus grandes fonctionnalités que Youtube. Reste que le site américain est encore le plus populaire avec plus de 60 millions de visites au mois de mai 2008.

Youtube est une plateforme de vidéos gratuite créée d'abord aux USA en février 2005 par trois ex-employés de Paypal : Chad Hurley, Steve Chen et Jawed Karim..

La société a son siège social à San Bruno en Californie.

Google rachète Youtube en octobre 2006 pour 1.65 milliards de dollars. La transaction s'est finalisée le 13 novembre 2006 et à partir de ce moment tout a changé. Des statistiques montrent qu'en 2009, 350 millions de personnes visitaient chaque mois ce site.

Youtube.fr et les versions européennes sont lancées en juin 2007 .

A quoi ca sert :

Youtube est un site web qui regorge de vidéos. C'est un site de partage, les utilisateurs peuvent envoyer, visualiser et partager des séquences vidéo.

Il utilise la technologie Adobe Flash pour l'affichage de toutes sortes de vidéos que ce soit des extraits de films, d'émissions de télé, des clips de musique ainsi que des vidéos amateurs.

La plupart des vidéos peuvent être visualisées par tous les internautes, mais en revanche uniquement les membres peuvent envoyer des vidéos de façon illimitée. Pour devenir membre, il suffit de vous inscrire, c'est simple et gratuit. Et en plus vous allez pouvoir poster des commentaires et même noter les vidéos.

Dans le moteur de recherche Youtube, **lorsque** vous choisissez un artiste, il y a un menu déroulant qui vous propose d'autres vidéos du même auteur et on vous propose d'autres vidéos en rapport avec celle visionnée qui se trouve à droite de l'écran dans une barre défilante grâce au titre et aux étiquettes.

Par la suite, le site a fixé une limite de 10 minutes pour la longueur des vidéos ce qui au final ne change à rien, car n'importe quel membre peut découper une vidéo longue en plusieurs parties de 10 minutes chacune.

Le magazine PC World a classé Youtube à la 9^e place de son classement des 10 meilleurs produits de l'année 2006.

3.3.1. Création d'un compte sur YouTube :

- Connectez-vous sur www.youtube.fr
- Cliquez sur Inscription en haut à droite de la page
- La boîte de dialogue suivante apparaît :
- Choisissez un Nom d'utilisateur disponible sous lequel vous serez désormais connu,
- Complétez vos données personnelles,
- Consultez les conditions d'utilisations et cliquez sur « J'accepte »

Une nouvelle boîte de dialogue apparaît pour relier votre compte YouTube à une adresse e-mail :

Si vous ne possédez pas de compte Google (avec une adresse en @gmail.com), complétez la partie droite de l'écran avec :

- votre adresse e-mail actuelle (elle n'apparaîtra pas sur YouTube).
- un mot de passe propre à votre compte YouTube que vous entrerez 2 fois,
- recopiez le texte de vérification affiché dans le champ juste au dessous

Créer mon compte

Rejoignez la plus grande communauté de partage de vidéos au monde !

- Recherchez et parcourez des millions de vidéos de la communauté et des partenaires
- Laissez des commentaires, donnez votre avis et publiez des commentaires vidéo pour vos vidéos préférées
- Envoyez vos vidéos pour les faire découvrir à des millions d'utilisateurs
- Enregistrez vos vidéos préférées pour les regarder et les partager ultérieurement

Nom d'utilisateur :
Votre nom d'utilisateur peut uniquement comporter des lettres A-Z ou des chiffres 0-9.
[Vérifier la disponibilité](#)

Lieu :

Date de naissance :

Sexe : Homme Femme

Permettre aux autres utilisateurs de trouver ma chaîne sur YouTube s'ils possèdent mon adresse e-mail

Conditions d'utilisation : Veuillez consulter les conditions d'utilisation Google et les conditions d'utilisation YouTube ci-dessous :

Conditions d'utilisation
1. Votre relation avec YouTube
1.1 L'utilisation du site Internet YouTube (le « Site Internet ») et de tous les produits, services et logiciels, y compris le lecteur vidéo

L'envoi de contenu dont vous n'êtes pas propriétaire constitue une infraction aux droits d'auteur et est répréhensible par la loi. Votre compte sera supprimé si vous envoyez un contenu qui ne vous appartient pas sur le site.

En cliquant sur "J'accepte", ci-dessous, vous acceptez les [conditions d'utilisation de YouTube](#), les [conditions d'utilisation de Google](#) et les [règles de confidentialité](#).

- Cliquez sur « Créer le nouveau compte et terminer »

Figure 3.43 : Page d'inscription de youtube

The image shows the YouTube login interface. At the top, there is a navigation bar with 'Accueil', 'Vidéos', 'Chaînes', and 'Communauté' tabs, a search bar with a 'Rechercher' button, and a yellow 'Envoyer une vidéo' button. Below this, a question asks 'Vous disposez déjà d'un compte Google ?'. Two paths are available:

- Oui. Ajouter YouTube à mon compte Google.** This path requires entering an email and a password, with a checkbox for 'Mémoriser mes informations sur cet ordinateur.' and a 'Connexion' button.
- Non. J'ai besoin d'un nouveau compte.** This path requires creating a new account by providing a current email address, a new password (with a security level indicator), and a confirmation of the password. A 'Créer le nouveau compte et terminer' button is at the bottom.

Figure 3.44 : Page de logging de youtube

3.3.2. Chargement de votre vidéo sur votre compte

Votre compte est créé. Votre fichier vidéo est prêt. Il est maintenant temps de le télécharger sur votre compte afin qu'il soit désormais visible par tous. Connectez-vous sur votre compte nouvellement créé, et, selon le site choisi, utilisez les étapes correspondantes ci-dessous.

Envoi d'une Vidéo sur YouTube

1. Cliquez sur le bouton "**Envoyer une vidéo**" dans le coin supérieur droit de n'importe quelle page YouTube.
2. Dans la fenêtre suivante, cliquez sur le bouton "**Parcourir**" et recherchez votre fichier vidéo dans le répertoire où vous l'avez créé. Sélectionnez le fichier que vous souhaitez envoyer.
3. Cliquez sur le bouton "**Envoyer une vidéo**".
4. La transmission de la vidéo va commencer. Elle peut prendre un temps variable selon la taille du fichier : pour 30 secondes de vidéo en MPEG-4, le fichier fait environ 3 Mo, et

nécessitera environ 1 minute avec un abonnement ADSL à 3Mo. N'oubliez pas que la vitesse



montante de votre ligne ADSL est en général le quart de la vitesse descendante.

Figure 3.45 : Page envoyer vidéo

5. Entrez toutes les informations utiles concernant la vidéo, notamment le titre, une description, des tags et une catégorie. Plus les informations sont nombreuses, plus les utilisateurs seront nombreux à consulter votre vidéo.

6. Pour que votre vidéo soit publique, choisissez « Partager votre vidéo avec le monde entier »

7. Cliquez sur le bouton "Enregister".

Votre vidéo est maintenant en ligne et les visiteurs de YouTube pourront la consulter.

You Tube Broadcast Yourself™
France | Français (0) mygra

Accueil Abonnements Vidéos Chaînes

Envoyer un fichier vidéo

✓ **medium.m4v** Réussite

[Ajouter un nom de vidéo, une description et modifier les paramètres de confidentialité](#) [Options d'intégration et de partage](#)

Titre
Henri Vennin - Assistance au pédalage par air comprimé

Description
Cette invention brevetée a été présentée par Henry Vennin au Salon International des Inventeurs de Genève en avril 2009. Il s'agit d'un procédé d'assistance au pédalage d'un cycle par utilisation de la force complémentaire apportée par un moteur fonctionnant à l'air comprimé fourni par une bouteille classique telle que celles utilisées pour la plongée sous-marine. Pour plus de détails, visitez le site www.mygraal-innovation.com

Tags
henri vennin, assistance pédalage, rosalie, attraction, air comprimé, invention, genève, inventeurs

Catégorie
Science et technologie

Droits d'accès

Partagez votre vidéo avec le monde entier (recommandé)
 Privée (vous pouvez partager la vidéo avec 25 personnes)

Enregistrer

Cliquez sur "Envoyer la vidéo" pour sélectionner et envoyer un fichier vidéo. **Envoyer la vidéo**

Nombre total de vidéos envoyées : 1 [Accéder à mes vidéos](#)

Figure 3.46 : Page de réussite d'envoyer vidéo

3.3. 3 L'utilisation du tags par Youtube :

Fondé en 2005 et racheté par Google en 2006, Youtube reste encore le site de référence pour le partage de vidéo. La recherche se fait avec des mots-clés, « Tags » utilisés par les utilisateurs, dans la barre de recherche. Depuis peu la barre de recherche affiche des suggestions de recherches (comme Google). Les utilisateurs qui téléchargent des vidéos les regroupent dans des chaînes (channels) et n'importe quel internaute peut s'y inscrire et ainsi être mis au courant des nouvelles vidéos de sa chaîne favorite.

Les internautes inscrits peuvent voter pour leurs vidéos favorites, les sauvegarder dans une liste de favoris, composer leurs « playlists » et commenter les vidéos regardées. Les vidéos peuvent aussi être téléchargées à partir d'un téléphone mobile.

Dans l'affichage vidéo, à droite de l'écran, s'affichent les vidéos du membre qui les met en ligne et encore des « related vidéos » avec les mêmes mots-clés. S'affichent aussi des vidéos sponsorisées par des professionnels et des publicités. Les vidéos peuvent être visionnées en plein écran mais leur qualité n'est pas toujours homogène, étant donné le taux de compression utilisé pour le téléchargement. Une fonctionnalité intéressante est d'utiliser un code, présent à droite des vidéos, qui les incorpore dans un cadre pour les afficher dans des blogs et des forums.

La page d'accueil regroupe les abonnements aux diverses chaînes, les suggestions de vidéos basées sur les mots-clés des favoris et les vidéos commerciales qui sont développés par des professionnels. Comme beaucoup d'autres sites, Youtube tend à regrouper plusieurs services sur la même interface. Ainsi, un service de courriel est disponible pour recevoir les messages d'autres utilisateurs et suivre leur activité. Le service permet aussi de suivre la trace des commentaires laissés sous une vidéo, permettant de voir qui a répondu. Un carnet de contacts permet d'organiser les différents contacts et de les classer dans des catégories avec des options de confidentialité diverses.

Plusieurs options sont disponibles pour protéger les informations personnelles des internautes inscrits, comme la possibilité de rendre des vidéos accessibles seulement pour les amis et la famille. Pour ce qui est du contenu, Youtube reste intéressant pour voir un extrait de concert, vidéo, ou une interview des médias, mais on y trouve également beaucoup de vidéos personnelles des utilisateurs qui parlent de leur vie sentimentale, leurs relations. Le fait de pouvoir interagir directement avec des gens qui parlent de problèmes bien réels a beaucoup d'attrait pour les internautes, mais il ressort qu'il y a beaucoup de contenu sans intérêt réel. Le problème le plus fréquemment rencontré est l'abus des mots-clés utilisés pour attirer le plus d'auditeurs pour une vidéo ou une chaîne. Ces mots-clés sont souvent sans rapport avec le contenu même de la vidéo.

Connectez-vous à You Tube et renseignez nom d'utilisateur et mot de passe et cliquez sur « Envoyer une vidéo ».

The screenshot shows the YouTube homepage with the following elements:

- Top navigation: [Inscription](#) | [Pense-bête \(0\)](#) | [Aide](#) | [Connexion](#) | [Site](#) |
- Main navigation: [Accueil](#) | [Vidéos](#) | [Chaînes](#) | [Communauté](#)
- Search bar: | [Vidéos](#) | [Rechercher](#) | [options avancées](#) | [Envoyer une vidéo](#)
- Section: **Se connecter à YouTube**
- Text: **Rejoignez la plus grande communauté de partage de vidéos au monde !**
Bénéficiez d'un accès complet avec votre compte YouTube :
- List of benefits:
 - Recherchez et parcourez des millions de vidéos de la communauté et des partenaires
 - Commentez et notez vos vidéos préférées
 - Envoyez vos vidéos pour les faire découvrir à des millions d'utilisateurs
 - Enregistrez vos vidéos préférées pour les regarder et les partager ultérieurement
- Login form:
 - YouTube logo
 - Nom d'utilisateur :
 - Mot de passe :
 - [Connexion](#)
 - [Nom d'utilisateur oublié](#) | [Mot de passe oublié](#) | [Aide](#)

Renseignez la fiche de votre vidéo, en choisissant de bons mots-clef (tags).

Envoyer une vidéo (étape 1 sur 2)

The screenshot shows the 'Envoyer une vidéo' form with the following fields and content:

- Titre :* :**
- Description :* :**
- Catégorie :* :**
- Tags :* :** Les tags sont des mots clés permettant d'aider les internautes à trouver votre vidéo. (séparés par des espaces)

Cliquez pour ajouter les tags suggérés : [Alternative](#), [Blues](#), [Classique](#), [Pays](#), [Électronique](#), [Folk](#), [Hip-hop](#), [Indépendant](#), [Jazz](#), [Pop](#), [R&B](#), [Rap](#), [Religieux](#), [Rock](#), [Soul](#), [Sans label](#), [Musiques du monde](#)
- Footnote: (* indique un champ obligatoire)

Datez votre enregistrement et localisez-vous (en draguant) le petit curseur rouge.

Recherchez votre fichier WMV sur le disque dur et « envoyez la vidéo »

Envoyer une vidéo (étape 2 sur 2)

The screenshot shows the second step of the video upload process:

- Text: **Sélectionnez une vidéo à envoyer :**
- File selection: [Parcourir...](#)
- Submit button: [Envoyer la vidéo](#)

Voilà, c'est terminé. Après quelques minutes de transfert et quelques autres minutes de traitement, votre vidéo sera « en ligne » sur YOU TUBE.

Toutefois, il faudra attendre quelques heures avant qu'elle n'apparaisse dans le moteur de recherche de You Tube et il vous faudra patienter quelques heures de plus avant d'avoir votre premier visiteur.

Chapitre VI

La navigation sur le Web

Introduction :

La navigation joue un rôle important en formant nos expériences sur le Web. Elle fournit l'accès aux données d'une manière dont améliore la compréhension, reflète la marque, et la prête à la crédibilité globale d'un site. Et finalement, navigation de Web et la capacité de trouver des informations pour avoir une incidence financière pour des déposataires.

La conception de navigation est une tâche qui n'est pas simplement limitée à choisir une ligne des boutons. Elle est beaucoup plus large, et, en même temps, plus subtile que celle. Le créateur de navigation coordonne des buts d'utilisateur avec des buts d'affaires. Ceci exige une compréhension de chacun, comme une profonde connaissance d'organisation de l'information, de mise en page, et de présentation de conception.

4.1 Qu'est-ce que la "navigation" :

La navigation (en anglais on dirait "browsing") convient bien à une recherche d'ordre général, un peu comme lorsque l'on parcourt les rayons d'une bibliothèque, une liste de nouvelles publications ou de nouvelles acquisitions, une table des matières ou un index... On ouvre une page sur un site pour parcourir et choisir ce qui nous intéresse. On trouve sur Internet différents répertoires qui présentent des sites selon divers classements : par ordre alphabétique, par ordre systématique, par sujet ou thème, par pays ou région.

Pour bien exploiter la fonction navigation, il suffit de connaître la notion d'adresse URL et les fonctionnalités de base de votre logiciel de navigation (*Netscape* ou *Internet Explorer*). On indique une adresse de site, on termine par la touche "retour" ou "enter". Une fois sur le site, on clique sur les liens hypertextes qui nous intéressent. Un lien hypertexte est signalé par du texte souligné ou par des boutons ou icônes. Lorsque l'on se trouve sur une page intéressante, on peut y mettre un signet (en anglais "bookmark") pour s'y référer facilement par la suite. La fonction "bookmark" ou "favoris" se trouve dans la barre de menus de votre logiciel et son contenu est gardé en mémoire sur le disque dur de votre ordinateur. Ce fichier peut être édité à votre guise, selon les fonctionnalités de votre logiciel de navigation.

4.2 L'hypertexte :

L'hypertexte est un concept qui, apparu en 1960 à l'initiative de Ted Nelson, ne se voit vulgariser que dans les années 1990. La propagation du terme en France est liée à l'apparition du World Wide Web, et aux nouvelles possibilités du traitement de l'information documentaire grâce aux ordinateurs.

« Un hypertexte est un ensemble de texte reliés par des liens. Ces liens, généralement représentés par un souligné et une couleur différente sont actifs. Naviguer sur le web est ainsi le strict équivalent de se déplacer dans un hypertexte. »

Le web est un système fondé sur l'utilisation de l'hypertexte, qui permet la recherche d'information dans Internet, l'accès à cette information et sa visualisation. L'hypertexte est avant tout caractérisé par sa non linéarité. En effet la plupart des définitions s'accordent sur ce point, par exemple l'Office de la Langue Française le définit ainsi :

« Présentation de l'information qui permet une lecture non linéaire grâce à la présence de liens sémantiques activables dans les documents.»

Ou encore :

« L'hypertexte permet un parcours non linéaire d'une œuvre (ou d'un document quelconque au demeurant) par des clics sur des zones actives qui masquent des liens avec d'autres éléments du texte.»

D'un point de vue lexicologique, « hypertexte » est un mot composé savant à base nominale formé en français du préfixe grec *hyper* et du nom commun *texte*. Le mot *texte* vient du latin *textus* qui signifie « le tissu, la trame ». Avec l'apparition de l'hypertexte on assiste à une remotivation du sens premier. En effet l'hypertexte, associé au *web* qui signifie toile, est remotivé dans la mesure où la toile, le tissu, et la trame appartiennent au même champ lexical.

Le préfixe *hyper* exprime la démesure, l'excès. Il signifie au « dessus, au delà ». Il faut donc imaginer un texte au delà du texte lui même, au delà d'un point de vue local : le texte se trouve après ou avant un autre texte mais aussi un texte démesuré, que l'on ne peut plus quantifier.

Ainsi la première partie de la définition d'hypertexte dans la deuxième édition du *Dictionnaire du multimédia*, de Jacques Notaise, Jean Barda, et Olivier Dussanter (AFNOR, 1996), explique que l'hypertexte est une fonction logicielle qui, dans un texte, associe, à certains mots ou groupe de mot, des programmes procurant des compléments d'information se situant au-delà du texte. A chaque instant le lecteur peut obtenir ces compléments en cliquant simplement sur le mot ou groupe de mots choisi. Certains logiciels hypertextes traitent ainsi tous les mots d'un texte en associant, par exemple, à chacun sa définition puisée dans un dictionnaire ou un glossaire.

Il est à signaler que les auteurs du *Dictionnaire de l'informatique de l'AFNOR* (Association Française de Normalisation), de 1997, ayant pour sous-titre *Le vocabulaire normalisé*, présentent l'hypertexte comme relevant de notre manière de penser, ce qui n'est pas sans implication historique. En effet, que l'hypertexte permet la réalisation de documents d'une manière qui imite la structuration humaine des idées, différentes de la structure linéaire des films, des livres et de la parole.

Pour conclure sur cette définition voici les principales caractéristiques de l'hypertexte de ce que l'on peut en retenir :

- il est lié au web et s'organise donc comme un réseau, un tissu ;
- ce tissu est composé de nœuds (la page) et de liens qui varient
- les liens peuvent être des images, des sons, des vidéos, appelés hypermédia ;
- l'hypertexte est à la fois un concept, un outil informatique, un produit ;
- il s'inscrit dans le domaine auquel il se réfère, c'est pourquoi les définitions, suivant qu'elles prennent telle ou telle dominante peuvent varier ;
- il s'oppose aussi à l'approche traditionnelle du livre dont il remet en cause certaines notions

4.2.1. Hypertexte : nœud et liens :

Cependant les sèmes de l'hypertexte ne se limitent pas à sa non linéarité, d'écriture et de lecture. Il est manifeste que la lecture plurilinéaire représente un topos récurrent, un trope inévitable dès qu'on évoque le sujet. Ce processus s'établit à l'aide de « liens ». Le lien permet de naviguer entre les documents, d'ailleurs c'est cette partie de l'hypertexte que le *Petit Robert* sur cédérom retient :

« Procédé permettant d'accéder aux fonctions ou informations liées à un mot affiché à l'écran, en cliquant simplement sur ce mot. »

En effet les notions de lien et de nœud sont inséparables de l'hypertexte dans la mesure où celui-ci est constitué de nœuds d'information connectables par des liens activables. Ainsi nous trouvons également :

« Hypertexte, nm, technique ou système qui permet dans une base de données documentaire de passer d'un document à un autre selon des chemins préétablis ou élaborés lors de la consultation. »

Et, « INFORM, système constitué d'un ensemble de textes associés par des liens permettant à l'utilisateur de naviguer de l'un à l'autre. » pour ce qui concerne le dictionnaire de la francophonie » Hachette, en ligne.

De plus une distinction est établie entre les différents types de liens : généralement sont distingués :

- le lien organisationnel et le lien sémantique,
- le lien interne et le lien externe à une page,
- le lien défini par l'auteur et le lien créé par le lecteur,
- le lien défini selon sa direction et dans son rapport avec le nœud
- etc.

Le lien, pour passer d'un document à un autre peut être de différentes natures : comme par exemple du texte : mot, groupe de mots, adresse url, images, sons.

4.2.2. Propriétés de l'hypertexte :

Le document hypertexte présente **plusieurs propriétés majeures**. Il est :

- **Non-linéaire**

Alors que la linéarité d'un livre ou d'un film impose à l'esprit un ordre figé et une séquentialité décidé par le concepteur, l'écriture et la lecture d'un hypertexte peuvent s'effectuer de manière non-linéaire; la lecture peut procéder par sauts d'un fragment à un autre, et l'écriture, par ajouts et retraites de nœuds et de liens, par constructions et déconstructions de réseaux et d'unités d'information. Cette non-linéarité permet, en outre, à un même document hypertexte de proposer et contenir une multiplicité de niveaux d'analyse, de niveaux de complexité et de niveaux de spécialisation.

- **Modulaire et combinatoire**

Chaque nœud d'informations du réseau est autonome. Les informations sont accessibles par différents chemins et dans différents contextes. Les blocs d'information peuvent être combinés, branchés/débranchés sur l'existant, à l'infini.

- **Interactif**

L'interactivité est une propriété qui découle des deux précédentes. Le lecteur interagit avec le texte. L'hypertexte laisse en effet la liberté au lecteur de décider de son cheminement dans le document en fonction de ses besoins, de ses envies, de ses intérêts, lui donne la liberté d'ignorer, d'effleurer, de découvrir, d'approfondir tel ou tel aspect. Le lecteur se construit ainsi, de lien en lien, de fragment en autre fragment de savoir, sa propre connaissance. L'internaute assigne du sens aux blocs d'information et de savoirs qu'il découvre ou prospecte à l'éclairage des contextes des documents qu'il choisit de visiter.

- **Immatériel et immédiat**

Son contenu est accessible et partageable indépendamment de son support physique. Il en découle que l'activation d'un lien met à disposition immédiate le contenu lié, sans recherche, ni manipulation de document (et donc sans perte de temps).

- **Continuellement modifiable**

Il peut être modifié ou actualisé à tout moment sans interrompre l'accès qui en est effectué dans les réseaux de documents dans lesquels il est maillé (lié).

- **Ouvert et infini, instable et dynamique**

Il n'est pas un objet clos et définitif, mais un contenu instable en relations et en mouvement.

4.2.3. Inconvénients de l'hypertexte :

L'hypertexte peut être aussi à l'origine de plusieurs problèmes :

- **la désorientation :**

Elle est liée notamment à la difficulté que peut rencontrer l'internaute pour comprendre l'organisation des connaissances et la stratégie d'apprentissage telles qu'elles ont été prévues par les concepteurs. L'internaute se retrouve alors face à un labyrinthe informationnel dont il ne parvient pas à mesurer l'abondance, ni à en appréhender la structure, ni donc à se faire une représentation du contenu. Il est désorienté dans le chemin qu'il parcourt et qui ne parvient pas à donner du sens à son exploration. Il se perd, par rapport à ses buts et à la façon de les atteindre.

- **la dispersion :**

La structure éclatée, insondable, non préhensible de l'hypertexte, le butinage non contrôlé de site en site, une exploration insuffisamment disciplinée peuvent faire naître un sentiment de dispersion en regard d'un objectif initial.

- **le coût cognitif**

La lecture hypertexte est cognitivement coûteuse car :

- le lecteur doit chercher et trouver l'information avant de pouvoir lui donner sens et l'interpréter.
- elle est complexifiée par la diversité des structures et par la pertinence des contenus à évaluer.

- **la surcharge cognitive**

Elle est liée au foisonnement d'informations.

4.2.4. Lien hypertexte (ou *hyperlien*) :

Élément activable (mot, expression, image, etc) d'une page permettant à l'internaute qui clique dessus d'accéder directement à une autre section (interne ou externe au document courant) sur laquelle pointe cet élément.

Pour donner la cohérence au document, les liens doivent être utiles, pertinents, compréhensibles et signifiants. On distingue :

- **les liens sémantiques** : à fonction cognitive, pédagogique,
- **les liens de navigation** : à fonction utilitaire.

4.3. Hypermédia :

Extension de l'hypertexte à des données multimédias.

Ensemble organisé d'informations, appartenant à plusieurs types de médias : textes, graphiques, images, animations, séquences vidéos, séquences audio (etc), pouvant être lus,

écoutés, vus. Comme l'hypertexte, il fonctionne sur le mode interactif et engendre de multiples parcours de lecture des informations.

Le terme multimédia désigne la possibilité d'utiliser des médias différents, des supports physiques différents : écrans, papier, images fixes ou animées, bandes sons... pour apporter l'information

L'Hypermédia fait référence à l'interaction "homme/connaissance". Il désigne un mode de gestion de l'information où les unités d'information sont connectées entre elles par des liens.

L'hypermédia, qui conjugue hypertexte et multimédia, est à l'heure actuelle le support le plus riche pour l'utilisateur, car il obéit à une structure réseau. Il se rattache aux *medias commutatifs* (type ordinateur, téléphone) par opposition aux *médias irradiants* (type presse, radio, télévision) qui diffusent l'information à partir d'un centre.

4.4. Hyperdocument :

Document électronique dans lequel la navigation n'est pas linéaire. L'hyperdocument peut ainsi désigner un hypertexte, un hypermédia mais aussi une base de données.

Un hyperdocument est un document *virtuel* qui n'est jamais globalement perceptible car son appréhension est conditionnée par le chemin de lecture *non-linéaire* que le lecteur choisira de parcourir. Cette propriété de l'hypertexte en fait un document *interactif* et à géométrie variable dans lequel le lecteur tient une place prépondérante.

Un système hypertextuel est « Un dispositif informatisé permettant l'interconnexion de documents de divers types, ceci non sur la base d'un modèle hiérarchique ou relationnel, mais par des mécanismes associatifs sous contrôle de l'utilisateur. (Bruillard, 1997)

L'utilisation d'hyperdocuments à l'école semble pouvoir accompagner le souci actuel d'une individualisation de l'enseignement :

- d'une part en imposant une relation individuelle à l'objet technique
- d'autre part en permettant un accès différencié à l'information

4.5. Navigation : espace de conception :

La navigation dans un espace d'informations est une tâche-utilisateur qu'il convient d'étudier en termes d'Interaction Homme-Machine (IHM). Dans [Dourish 94], les auteurs proposent trois modèles de navigation représentés à la Figure 4.1 : spatial, sémantique et social.

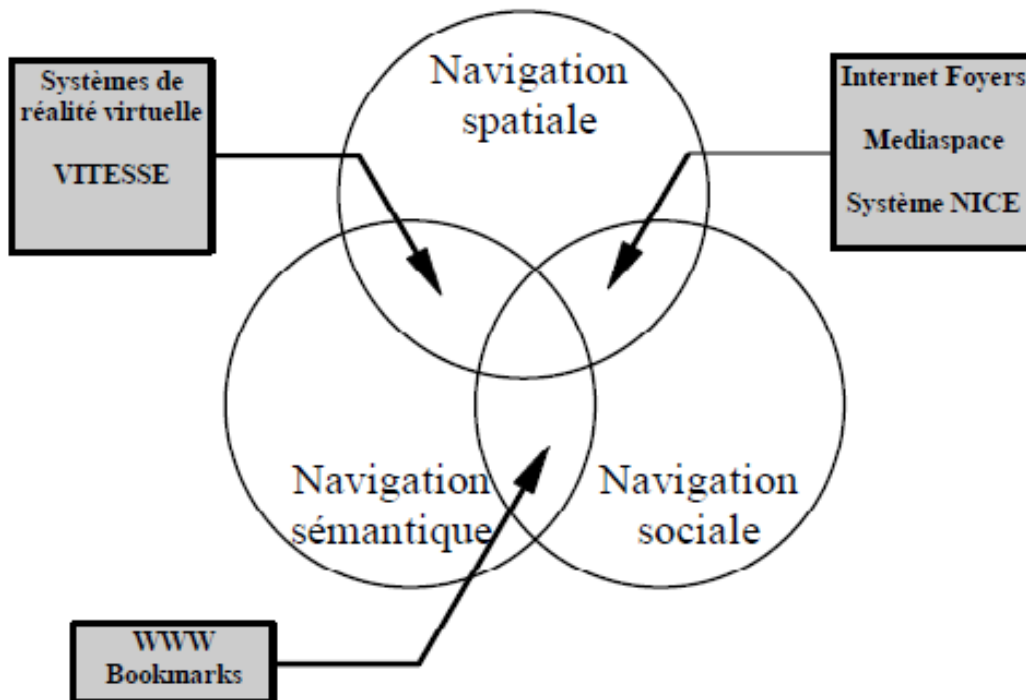


Figure 4.1 : Les trois modèles de navigation.

4.5.1. Navigation spatiale :

Le modèle spatial repose sur l'analogie avec le monde réel et en particulier avec nos connaissances sur l'espace (notion de proximité, de perpendicularité, d'alignement, de taille, etc.). Ce modèle de la navigation est autant utilisé dans les systèmes de réalité virtuelle que dans de nombreux systèmes d'informations. Les aspects dynamiques du modèle spatial sont aussi intéressants car la navigation implique des objectifs à atteindre, comme trouver une information ou un ensemble d'informations.

De plus la dimension temporelle joue un rôle primordial [Mackinlay 91]. Le temps et l'espace sont deux ingrédients en intime relation qui caractérisent la navigation. La question

"comment aller à ..." identifie outre l'aspect spatial (la direction et le chemin à emprunter) un autre paramètre important : le moyen de locomotion. La Figure 4.2 organise selon deux dimensions les caractéristiques de la navigation [Vernier 97b].

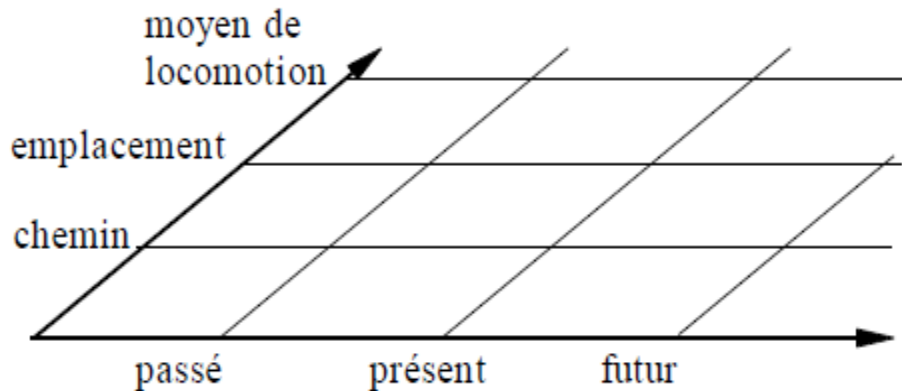


Figure 4.2 : Les caractéristiques de la navigation.

La navigation purement spatiale peut être illustrée par assez peu de systèmes. Par exemple, les systèmes d'aide à la navigation permettent de naviguer dans l'espace afin de préparer l'itinéraire d'un voyage. Cette activité fait intervenir les chemins, les emplacements et les moyens de locomotion (train, voiture, etc.). D'un autre côté, l'utilisateur manipule les relations temporelles entre les emplacements et les chemins afin d'organiser son voyage.

4.5.2. Navigation sémantique :

Ce modèle de navigation décrit les comportements lorsque l'utilisateur se déplace dans l'espace d'informations en fonction d'attributs propres aux informations qui lui sont présentés (valeurs, similitude, etc.). Ce type de navigation permet l'accomplissement de la quasi-totalité des tâches de l'utilisateur. Sa mise oeuvre est donc primordiale. Par exemple, les systèmes hypertextes utilisent une navigation sémantique (déplacements le long des liens hypertextes) qui n'exploite pas les propriétés spatiales de l'information. Les utilisateurs du World Wide Web (WWW ou Web) naviguent selon ce modèle de navigation. Dans les butineurs du Web, le passage d'un document à un autre s'opère par un clic de la souris sur un mot ou une image.

L'emplacement spatial du mot ou de l'image n'a pas d'incidence sur la destination du lien.

4.5.3. Navigation sociale :

La navigation sociale est la troisième catégorie de navigation. Ce modèle de navigation repose sur l'utilisation d'informations à propos d'autres utilisateurs pour se déplacer (c.-à.-d. changer son centre d'intérêt). Cette forme de navigation suppose que l'on ait plusieurs utilisateurs qui partagent un même espace d'informations. Les comportements types de ce modèle sont les regroupements de personnes intéressées par un même sujet.

La navigation se fait en dehors du site. La responsabilité de l'éditeur du site est de s'assurer que les utilisateurs de plates-formes sociales sont compatibles avec le site. Proposez sur votre site des flux RSS, créez des widgets, intégrez des pictogrammes de type , delicious etc. Soyez actif dans ces communautés.

Naviguer socialement, c'est cliquer sur des liens qui sont proposés par la communauté. On parle de communauté, car ces liens sont soumis sur une plate-forme sociale par un individu. D'autres utilisateurs de cette plate-forme votent pour ce contenu. Plus un contenu sera recommandé, plus il sera mis en avant. Formule mathématique exponentielle, des déclinaisons de modèle de navigation sociale sont présentées sur la majorité des sites communautaires web 2.0 : YouTube, delicious... Des artistes percent et sont contactés par de grands labels de musique pour leur popularité sur le web.

La métaphore sociale de navigation (Dourish et Chalmers 1994 ; Erickson 1996 ; Dieberger 1997) est assez nouveau et a été utilisé pour saisir certains aspects de la navigation. Les définitions varient de « permuter la flèche indicatrice dans le WWW » à la navigation socialement améliorée ».

La présomption utilisée en ce document est que la navigation sociale est navigation. En fait, à la suite de la discussion ci-dessus sur la navigation, la seule chose qui sépare navigation et la navigation sociale sont les outils. Suite à ceci, pour comprendre la navigation sociale que nous devons étendre nos outils de navigation pour incorporer les éléments sociaux. Par conséquent quand nous dirigeons socialement il y a un plus d'outil disponible à nous, à savoir un agent ou un groupe des agents qui peuvent nous aider à diriger dans un espace de l'information. En outre d'autres outils non-social sont encore à la disposition d'un utilisateur,

par exemple si une personne est détruite dans une ville qu'ils peuvent demander à une autre personne (l'outil social) de préciser leur destination sur une carte. Les ayant fait ainsi peuvent alors employer la carte (l'outil non-social) pour trouver leur destination. De même, dans un espace de l'information virtuelle un utilisateur peut demander un agent une destination spécifique et puis employer des bornes limites pour diriger à elle. Peut-être il est plus approprié de penser à la navigation sociale en tant que navigation socialement améliorée. Pour la clarté je diviserai la navigation sociale en deux catégories :

- Navigation sociale directe
- Navigation sociale indirecte

L'agent a plusieurs caractéristiques et propriétés. Premièrement, l'agent peut être artificiel ou humain. Deuxièmement, un utilisateur peut communiquer directement ou indirectement avec l'agent à diriger dans l'espace de l'information. Un utilisateur communique indirectement quand il n'y a aucune transmission mutuelle entre l'agent et l'utilisateur. Dans un musée, par exemple, tour guidé serait classifié en tant que navigation sociale indirecte si le guide (agent) prend le spectateur (utilisateur) autour le musée, et là n'est rien que le spectateur peut faire pour modifier l'artère ou le pas de l'excursion.

Cependant, il n'y a aucune ligne nette entre indirect et la navigation sociale directe, par exemple, l'excursion guidée de musée peut également être considéré en tant que navigation sociale directe si le spectateur (utilisateur) a la capacité d'influencer l'excursion. Troisièmement, selon quel type de l'espace de l'information et de tâche de navigation, les agents peut être utile dans différentes voies. Dans un espace complexe de l'information tel qu'un équipement de traitement de texte, un utilisateur a souvent des idées vagues sur où elles veulent disparaître. Suite à ceci si elles avaient la capacité de communiquer directement avec l'agent, l'agent pourrait poser des questions et essayer de clarifier le but ou la destination de l'utilisateur.

- **Navigation sociale directe**

La navigation sociale directe est caractérisée par le fait qu'il y a mutuel-transmission entre l'utilisateur et l'agent. Un utilisateur peut poser des questions comme « où suis je ? » ou, « où peux je trouve l'emplacement X ? » L'agent répond que l'utilisateur et demande peut-être d'une manière primordiale à l'utilisateur de clarifier leurs questions. Un agent peut aider à

clarifier les buts de l'utilisateur ou même à les changer. Quand les utilisateurs sont incertains d'où ils veulent aller l'agent peut les supporter en formulant une destination.

Il y a trois formes de navigation sociale directe ; premièrement, un utilisateur peut communiquer avec un agent artificiel ; deuxièmement, un utilisateur communique avec un agent humain ; et en conclusion, un utilisateur communique avec un agent qui utilise alternativement d'autres agents. Chaque situation a de certains avantages et remours.

- **Navigation sociale indirecte**

Comme indiqué plus tôt dans la navigation sociale indirecte il n'y a aucune transmission bidirectionnelle entre l'utilisateur et l'agent. Certaines des techniques présentées ci-dessous peuvent sembler peu des forcées pour l'appeler navigation sociale comme nous l'avons définie, c.-à-d. un utilisateur communique avec un agent ou des agents à diriger un espace de l'information.

4.6. Navigation par tags :

Les systèmes de navigations par tag sont courants dans les blogs où ils complètent le classement des articles en catégories. Ils sont également bien pratiques sur des sites ou applications intégrant des volumes de données importants comme Flickr, Delicious.

4.6.1. Les tags, un outil de description et de navigation :

les tags permettent de décrire un contenu en fonction de ses caractéristiques. Ils se matérialisent par un lien hypertexte, contenant un mot-clé, et pointent soit vers une page, proposant une définition dudit mot-clé, soit, plus souvent, vers une page regroupant un ensemble de contenus décrits par ledit mot-clé.

La page vers laquelle pointe le lien peut aussi bien être interne au site qu'externe. On pourra ainsi pointer vers la page de résultats d'un outil de recherche prenant les tags en compte, comme Delicious.

La différence entre les deux modes d'utilisation, description ou agrégation est fondamentale.

Dans le premier cas, le tag est utilisé comme un élément descriptif du contenu, et peut contenir un terme que je ne rencontrerai pas dans ce dernier. Par exemple, je peux tout à fait écrire un article sur l'encéphalopathie spongiforme bovine sans utiliser une seule fois le mot "vache", tout en plaçant ce dernier dans mes tags, par exemple avec "vache" et "vache folle".

Dans le second cas, et c'est son utilisation la plus fréquente, le tag devient un mode de navigation à l'intérieur du site, de manière transverse là où les catégories sont un mode de navigation vertical. Je peux ainsi utiliser le tag "google" dans des billets classés dans les catégories "ergonomie", "développement", "humeur" sans qu'ils aient d'autre rapport entre eux que de parler de Google.

4.6.2. Utilisation des tags :

Les tags sont généralement utilisés comme navigation alternative et regroupent des documents en fonction d'expressions auxquelles ils peuvent être rattachés. Sur un des sites qui seront étudiés ci dessous, j'ai pris le parti de faire une navigation entièrement par tags. L'utilisation des tags a des avantages :

- On peut choisir librement les expressions couvertes. C'est beaucoup plus moins contraignant que de devoir utiliser les libellés d'un système de navigation classique.
- On peut avoir un grand nombre de liens vers les pages de tags, par exemple s'ils sont présentés dans un nuage présent sur toutes les pages du site.

4.6.3. L'intérêt des nuages de tags pour simplifier la navigation :

- Ils ne font pas appel à une notion trop structurée de l'information. Des études ont démontré que moins les outils sont structurés et plus les utilisateurs s'en emparent facilement et rapidement (ex : les wiki, les blogs, le Chat,...).

- La gestion de la taille ou de la couleur des mot-clés dans un nuage permet de rendre immédiatement visible l'importance relative d'un thème dans un ensemble de documents.

- Si les auteurs peuvent mettre des tags (mot-clés) sur les documents, il est aussi possible de permettre aux lecteurs d'y mettre leur propre mot-clés, afin de catégoriser les contenus selon leur propre schéma de pensée. Le site del.icio.us en est un exemple caractéristique.

- L'information structurée fait souvent appel, dans les outils à notre disposition, à une notion d'arborescence Cette arborescence est souvent le reflet de notre organisation mentale et de notre capacité à structurer nos données. Cependant, cette organisation peut-être logique pour nous et totalement incompréhensible pour d'autres. Ainsi voit-on des intranets ou des sites web dans lesquels rechercher du contenu se révèle être un véritable challenge. Utiliser des Tags (Mot-clés) présentés sous forme de nuage permet donc d'éviter au lecteur de devoir comprendre l'organisation documentaire des auteurs.

- L'utilisation des Tags permet de trouver un document sans utiliser un moteur de recherche de manière explicite. De cette manière est-on sûr de ne pas rechercher un mot-clé qui n'existe pas, ce qui est souvent une expérience désagréable. L'affinage de la sélection reste possible en sélectionnant tag après tag, comme cela peut fonctionner lorsque l'on utilise les filtres dynamiques des moteurs de recherche avancés. On trouve ce type d'affinage par exemple sur les sites techniques de Microsoft.

- Autre intérêt de la gestion documentaire par tags : les auteurs ne sont pas obligés de créer une arborescence documentaire. Cela permet donc de réorganiser la documentation de manière simple (par remplacement de tag) au besoin. Et ceci est loin d'être neutre lors des réorganisations internes...

Utiliser les tags et les nuages de tags permettra de simplifier la gestion documentaire tant pour les auteurs que pour les lecteurs. La représentation graphique des nuages de tags pourra varier en fonction des besoins (arborescence logique, mise en exergue,...).

4.6.2. Résultats en termes de référencement :

Nous avons comparé les résultats de l'utilisation des tags sur deux sites :

- Site 1 : a navigation principale se fait avec des tags
- Site 2 : sa navigation est classique et où les tags ne servent qu'à couvrir des expressions non présentes dans le reste de l'arborescence

Résultat sur un mois glissant de statistiques :

	Site 1	Site 2
Part des pages vues	17,69%	2,05%
Part des pages d'entrée	50,64%	2,78%

Tableau4.1 : résultats de l'utilisation des tags sur deux sites

Sur le premier site, les performances des pages de tags sont très intéressantes : elles ne constituent que 17% des pages vues, mais représentent plus de la moitié des entrées sur le site. L'avantage de l'utilisation exclusive des tags est bien visible : ces pages ont des contenus pertinents et peuvent recevoir des liens externes qui améliorent leurs performances. Sur site 2, les performances sont beaucoup plus modestes, malgré la présence d'un nuage de tags sur toutes les pages. C'est probablement dû au choix des expressions et au bon comportement de l'arborescence classique qui draine beaucoup plus de trafic.

Au final, les pages de tags peuvent constituer un plus intéressant sur un site, mais en dehors d'une utilisation exclusive, il ne faut pas trop en attendre.

4.7. Navigation et Social Tagging des systèmes :

Les gens tendent naturellement à classer des choses par catégorie. Il est comment nous semblons raisonnable du tiroir de monde-un pour les chaussettes, des autres pour des chemises ; nourriture dans cette cuisine compartiment, paraboloïdes dans celui-là. Le groupement et l'écriture de labels nous aide à trouver des choses plus tard.

Dans le monde numérique, nous créons les systèmes personnels de l'organisation avec des fichiers informatiques et des chemises. Certaines caractéristiques intrinsèques d'information, telles que le type de fichier ou la date, aident en te permettant de trier les emails par le temps reçu ou de rechercher seulement le PDFs, par exemple. Mais vous également créez des chemises et donnez des noms de fichiers.

Toutes ces recherches et organisation se fondent sur des méta-données, qui est généralement définie en tant que « données au sujet des données. Les » méta données sont les étiquettes employées pour décrire des objets : documents, livres, photographies, MP3s, pages Web, et ainsi de suite. Ils nous donnent des traitements à l'encavateur sur en organisant l'information. Un ordinateur ne sait pas qu'une photo sur votre disque dur est de Tour Eiffel pris en votre voyage à Paris en 2005 à moins que vous le disiez ainsi. En attachant des méta-données à l'image, la conclusion d'elle devient plus facile.

Au sens le plus large du mot, tagging, n'est pas une nouvelle activité. Bibliothécaires et sélecteurs « tagging » des livres avec les vedettes matières et les mots-clés pendant des années. Les tags sont simplement des méta-données que nous nous appliquons aux ressources et aux objets sur le Web ainsi nous pouvons les trouver plus tard. Cependant, la voie dont les tagging obtiennent fait dans les systèmes ouverts et de collaboration sur le Web using de tels services comme del.icio.us et Flickr est seul et innovateur.

Mais pour être utile, un système de tagging doit également être navigable. La discussion ici se concentre donc sur diriger les systèmes de tagging sociaux.

4.8. Systèmes de Tagging :

Il y a beaucoup de types tagging des systèmes, mais ils tous ont trois éléments clé :

Ressources : L'objet étant tagging, comme un page Web ou une photo

Utilisateurs : Les membres d'un système de tagging social qui appliquent des étiquettes aux ressources

Tags : Les termes employés pour décrire une ressource

Pratiquement chaque ressource sur le Web peut être tags. Une distinction importante affaires dans la façon dont des tagging sont créées, mises à jour, et utilisées avec la source d'objet à taggingr. L'origine de la ressource à tagging affectera qui l'étiquettera et comment il a tags. Les ressources à étiqueter peuvent être :

Fourni par des participants, tels que les photos téléchargées sur Flickr ou les vidéos sur YouTube

Fourni par le site Web ou le système de tagging, tel que des descriptions d'artiste et d'album sur Last.fm

Ouvrir les ressources existant sur le Web, tel que des pages Web bookmarking sur del.icio.us

Noter également que le processus pour tagging diffère nécessairement basé sur la ressource en question. S'applique tagging à un page Web public, par exemple, différent d'appliquer des étags à une photo ou à un vidéo que vous téléchargez.

Avec le page Web, le système peut identifier la ressource par son URL et le point à d'autres personnes de tags déjà utilisé. Le texte et le titre de la page peuvent également être employés pour suggérer des tags. Avec une photo ou une vidéo il n'y a aucune méta-donnée descriptive de préexistence et le système ne peut pas identifier cette ressource jusqu'à ce que vous l'expliquiez dans les mots.

Les tags peuvent également être l'une ou l'autre privée, de sorte que seulement l'utilisateur puisse les voir et utiliser, ou public, de sorte que le groupe entier puisse les utiliser. Gmail, par exemple, te permet de tagging les emails, mais ceux-ci jamais ne sont autrement vus par n'importe qui. Quelques services permettent à des utilisateurs de décider si les tags devraient être privées ou publiques, comme del.icio.us et Flickr.

Chapitre V

Réalisation

Les Tags représentent un nouveau modèle de structuration des données sur le web qui se fonde sur l'indexation libre de mots-clés par les internautes et offrent de nouveaux supports aux formes préexistantes de navigation sociale. On a proposé une typologie d'usages des systèmes de tagging. Cette activité de catégorisation, apparaissant comme une pratique individuelle, amène pourtant les utilisateurs à entrer dans de multiples formes de coopération, allant de coopérations implicites, rendues possible par l'agrégation des metadonnées par le système, à des formes de coopérations plus coordonnées au sein de collectifs d'utilisateurs.

5.1. Les motivations des langages :

Comme on l'a mentionné précédemment, on a choisie le paradigme OOP et plus exactement le langage python version 2.6 et l'environnement de Django version 1.1.1 comme des outils de développement pour les raisons suivantes :

5.1.1. Python :

Présentation du langage :

Python est un langage portable, dynamique, extensible, gratuit, qui permet (sans l'imposer) une approche modulaire et orientée objet de la programmation. Python est développé depuis 1989 par Guido van Rossum et de nombreux contributeurs bénévoles.

Caractéristiques du langage :

- **Python est portable**, non seulement sur les différentes variantes d'*Unix*, mais aussi sur les OS propriétaires : *MacOS*, *BeOS*, *NeXTStep*, *MS-DOS* et les différentes variantes de *Windows*. Un nouveau compilateur, baptisé JPython, est écrit en Java et génère du *bytecode* Java.

- Python est gratuit, mais on peut l'utiliser sans restriction dans des projets commerciaux.

- Python convient aussi bien à des scripts d'une dizaine de lignes qu'à des projets complexes de plusieurs dizaines de milliers de lignes.

- La syntaxe de Python est très simple et, combinée à des types de données évolués (listes, dictionnaires,...), conduit à des programmes à la fois très compacts et très lisibles. A

fonctionnalités égales, un programme Python (abondamment commenté et présenté selon les canons standards) est souvent de 3 à 5 fois plus court qu'un programme C ou C++ (ou même Java) équivalent, ce qui représente en général un temps de développement de 5 à 10 fois plus court et une facilité de maintenance largement accrue.

- Python gère ses ressources (mémoire, descripteurs de fichiers...) sans intervention du programmeur, par un mécanisme de comptage de références (proche, mais différent, d'un *garbage collector*).

- Il n'y a pas de pointeurs explicites en Python.

- Python est (optionnellement) multithreads.

- Python est orienté-objet. Il supporte l'héritage multiple et la surcharge des opérateurs. Dans son modèle objets, et en reprenant la terminologie de C++, toutes les méthodes sont virtuelles.

- Python intègre, comme Java ou les versions récentes de C++, un système d'exceptions, qui permettent de simplifier considérablement la gestion des erreurs.

- Python est dynamique (l'interpréteur peut évaluer des chaînes de caractères représentant des expressions ou des instructions Python), orthogonal (un petit nombre de concepts suffit à engendrer des constructions très riches), réflexif (il supporte la *méta programmation*, par exemple la capacité pour un objet de se rajouter ou de s'enlever des attributs ou des méthodes, ou même de changer de classe en cours d'exécution) et introspectif (un grand nombre d'outils de développement, comme le *debugger* ou le *profil*, sont implantés en Python lui-même).

- Comme *Scheme* ou *SmallTalk*, Python est dynamiquement typé. Tout objet manipulable par le programmeur possède un type bien défini à l'exécution, qui n'a pas besoin d'être déclaré à l'avance.

- Python possède actuellement deux implémentations. L'une, interprétée, dans laquelle les programmes Python sont compilés en instructions portables, puis exécutés par une machine virtuelle (comme pour Java, avec une différence importante: Java étant statiquement typé, il est beaucoup plus facile d'accélérer l'exécution d'un programme Java que d'un programme Python). L'autre génère directement du *bytecode* Java.

- Python est extensible : comme *Tcl* ou *Guile*, on peut facilement l'interfacer avec des bibliothèques C existantes. On peut aussi s'en servir comme d'un langage d'extension pour des systèmes logiciels complexes.

- La bibliothèque standard de Python, et les paquetages contribués, donnent accès à une grande variété de services : chaînes de caractères et expressions régulières, services UNIX standards (fichiers, *pipes*, signaux, sockets, threads...), protocoles Internet (Web, News, FTP, CGI, HTML...), persistance et bases de données (Gadfly, Sqlite on a utilisé cette dernière Dans notre implémentation de la BD sous python), interfaces graphiques.

- Python est un langage qui continue à évoluer, soutenu par une communauté d'utilisateurs enthousiastes et responsables, dont la plupart sont des supporters du logiciel libre. Parallèlement à l'interpréteur principal, écrit en C et maintenu par le créateur du langage, un deuxième interpréteur, écrit en Java, est en cours de développement.

- Enfin, Python est un langage de choix pour traiter le XML.

Pour finir nous pouvons dire que python possède toutes les caractéristiques et la puissance d'un langage pouvant créer des Agents en général et dans le domaine des applications web intelligent puisque il possède une grande puissance de traitement de texte et plusieurs bibliothèque sont développé dans ce domaine tels que : beautifulsoup pour le parsing des document html et feedparser pour le parsing des document XML et urllib,urllib2 pour ouvrir les URL et digérer l'authentification, les redirections, les cookies et autre choses Simplejson pour accéder au service de recherche de Google (Ajax search API) .

5.1.2. Django

Définition :

Django est un framework web, écrit en langage Python, et dédié à la publication sur Internet, propose une architecture MVC pour découpler le Modèle, la Vue et le Contrôleur. Un framework consiste en un ensemble de bibliothèques permettant le développement rapide d'applications abouties. Ces briques logicielles sont organisées pour être utilisées en interaction les unes avec les autres.

Pourquoi nous avons choisi Django ?

Pour faciliter notre travail dans les implémentations de l'interface graphique et création de la base de données on a intégré le python dans Django.

5.2. L'installation l'environnement de travail :

5.2.1. Installer python sous Windows :

- Vous aurez besoin d'aller à <http://www.python.org/download/> et télécharger la dernière version.

- Ensuite, double-cliquez sur le fichier.Exe et suivez les instructions d'installation.

- Une fois l'installation a été fait, nous devons ajouter le répertoire Python pour le chemin du système afin que nous puissions accéder à Python en utilisant l'invite de commande.

Maintenant que Python est installé et opérationnel, nous sommes presque prêts. Ensuite, nous allons installer Django et assurez-vous que nous avons un système de base de données.

5.2.2. Installer Django sur Windows :

- Pour télécharger Django, à la tête <http://www.djangoproject.com/download/>, et récupérer la dernière version officielle.

- Après avoir téléchargé l'archive de Django, il extrait le disque C.

- C: \> cd c: \ Django-x.xx

- C: python \> Django-x.xx setup.py install

- La dernière étape est de copier le fichier django-admin.py de Django-x.xx \ bin django \ à quelque part dans votre chemin d'accès système, tels que c: \ windows ou le dossier où vous avez installé Python.

- Une fois que cela a été fait, vous pouvez supprimer le dossier C: \ Django-x.xx, parce qu'il n'est plus nécessaire.

- C'est ça. Pour tester votre installation, ouvrez une invite de commande et tapez la commande suivante: c: \ django>-admin.py - version

Si vous voyez la version actuelle de Django imprimé sur l'écran, puis tout est réglé. Installation d'un système de base de données. Si vous avez Python 2.5 ou supérieur, vous n'aurez pas besoin d'installer quoi que ce soit, étant donné que Python 2.5 est livré avec le système de gestion de base de données SQLite contenues dans un module nommé sqlite3.

5.3. Description de l'application « Django Bookmarks » (interfaces et fonctionnalités de l'application) :

5.3.1. L'interface de l'application :

Après le lancement du serveur l'utilisateur doit être devant l'interface illustrée dans la figure suivante

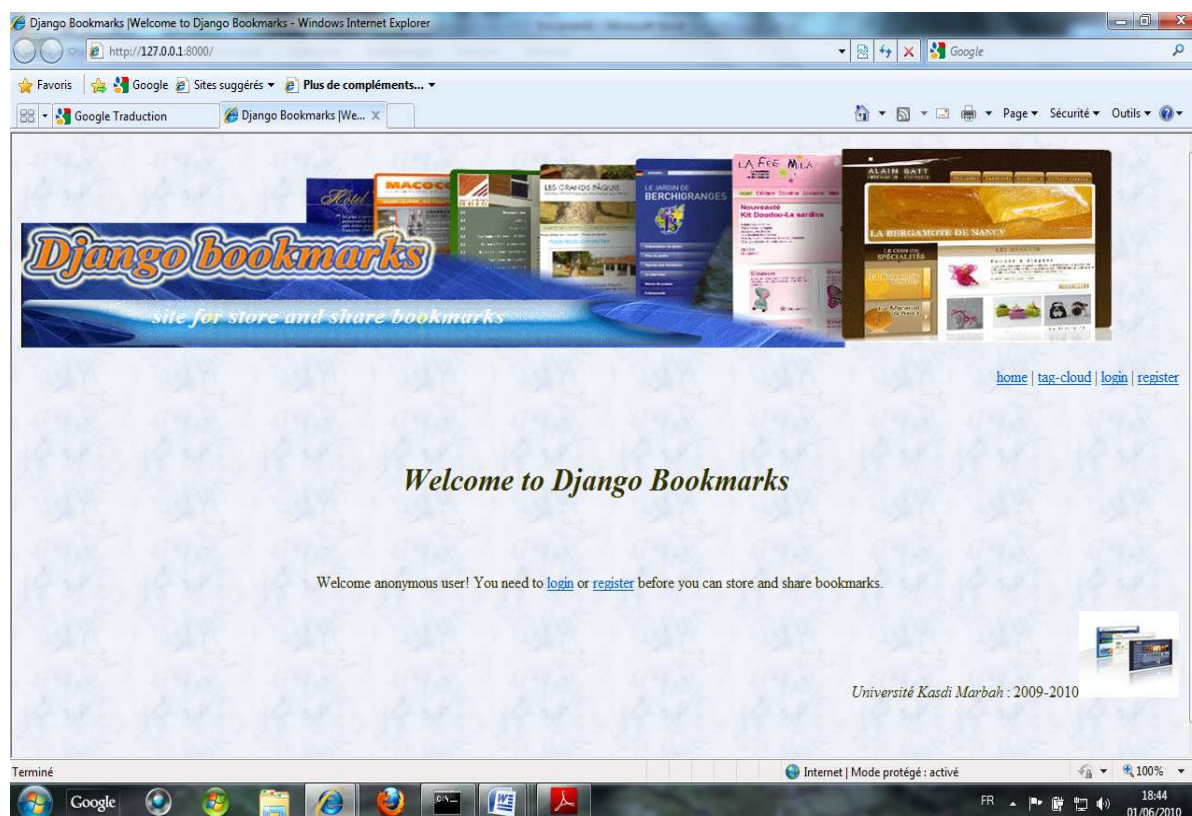


Figure 5.1 : La page principale de site Django Bookmarks

Au premier contact avec l'application, l'utilisateur doit saisir son username et son mot de passe pour faire un login ou une identification au système s'il est précédemment enregistré. (Figure 5.2)

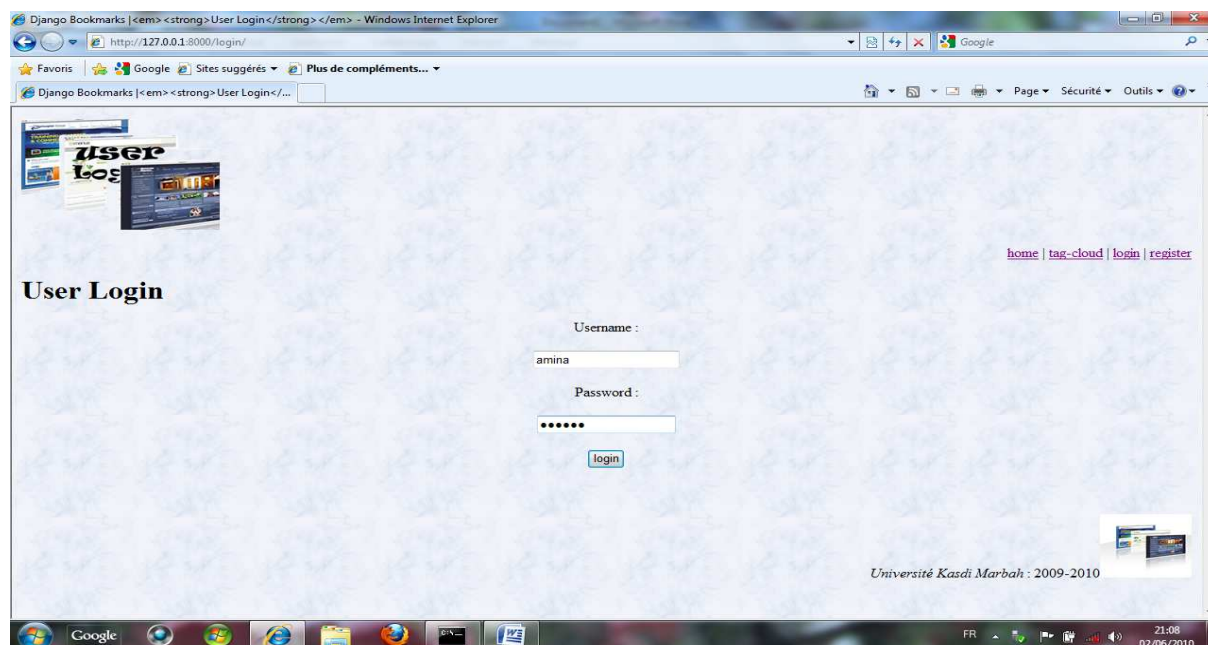


Figure 5.2 : Interface de l'accès au système et d'identification de l'utilisateur

S'il n'est pas il doit choisir register et doit remplir la page d'inscription comme il est montré dans la figure suivante.

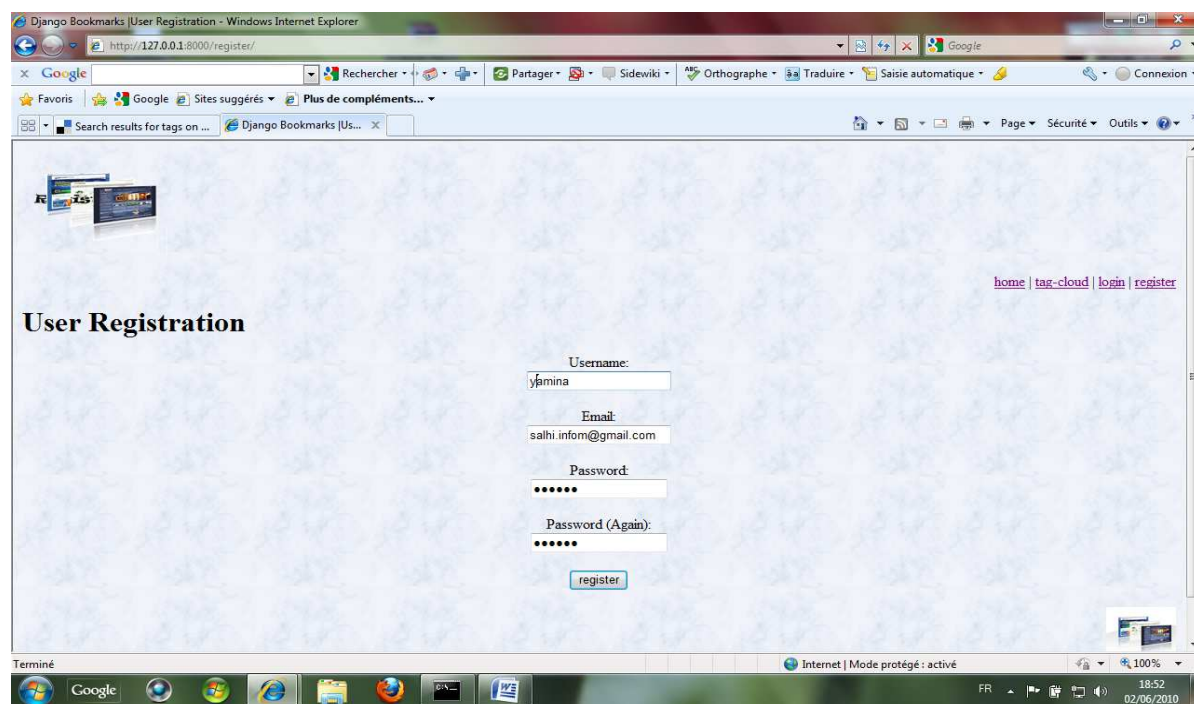


Figure 5.3 : Interface d'inscription du système

Et voici l'interface qui démontre le succès du processus d'enregistrement

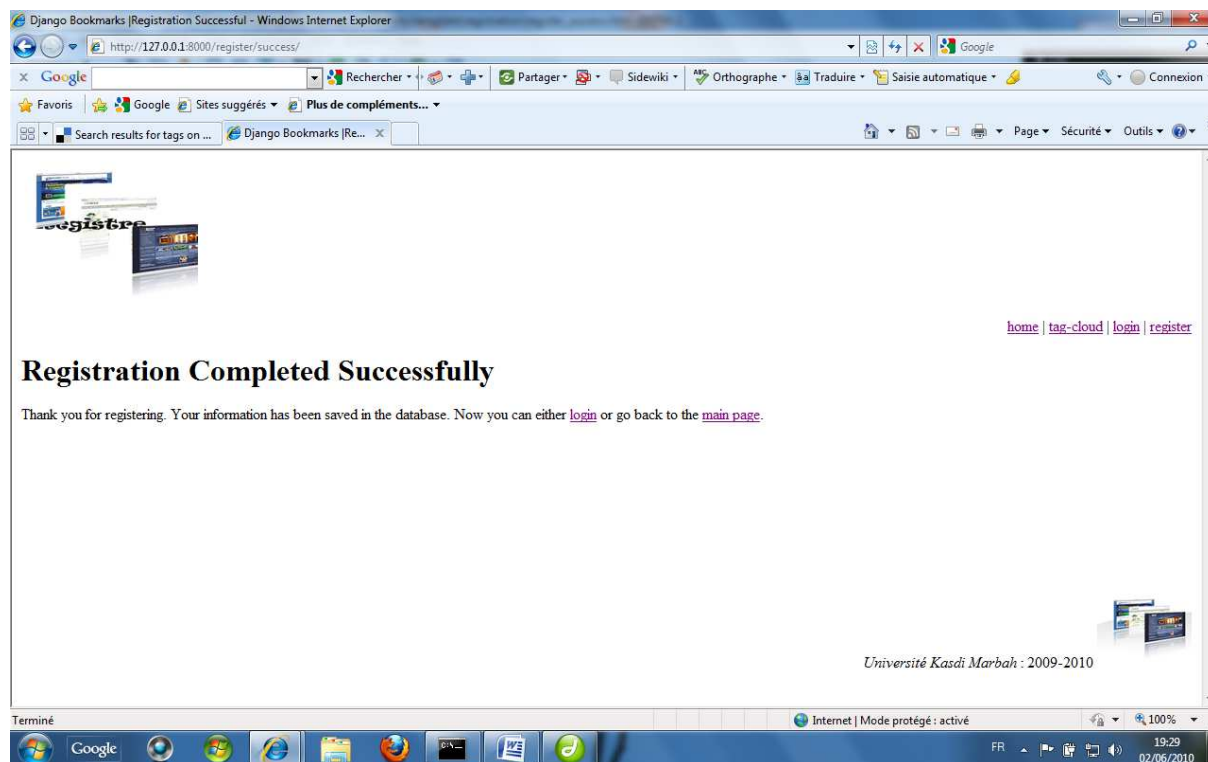


Figure 5.4 : Interface due succès du d'enregistrement

Après l'identification le système affiche l'interface principale de utilisateur qui permet a lui de stocker et partager ses bookmarks

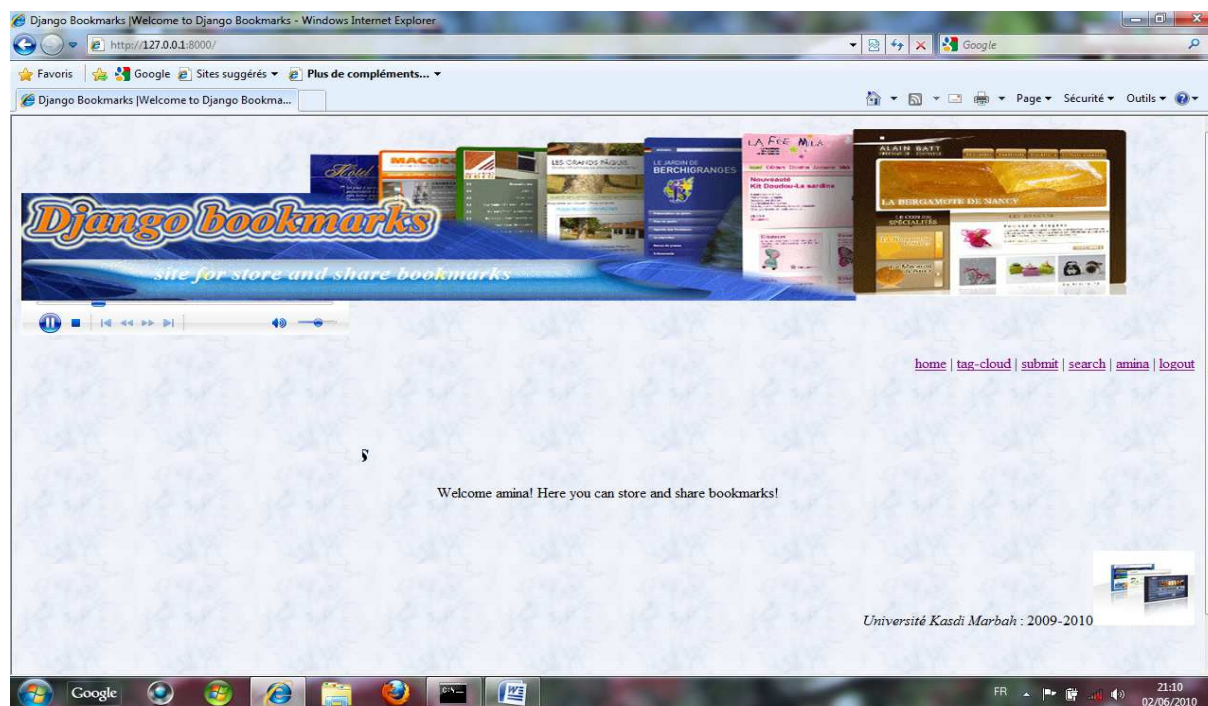


Figure 5.5 : Interface utilisateur

5.3.2. Sauvegarder les bookmarks:

L'utilisateur peut sauvegarder les sites préférés par l'identification de leur lien (URL), et leur (titre) et le groupe de Tags auquel il appartient ; La figure suivante illustre cette phase.

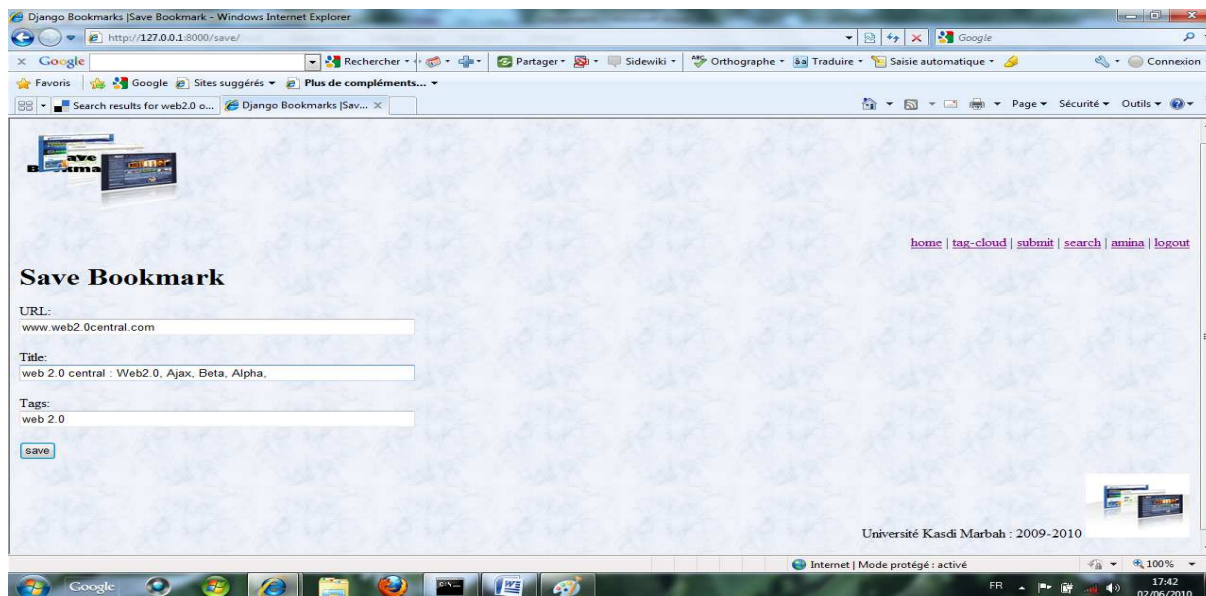


Figure 5.6 : Interface de Sauvegarder les bookmarks

5.3.3. La recherche des bookmarks

Pour effectuer une recherche d'un bookmark, l'utilisateur doit d'abord cliquer sur Search puis saisir le terme a été recherché et le système doit afficher la liste de tous les sites disponibles sur ce terme comme suit :

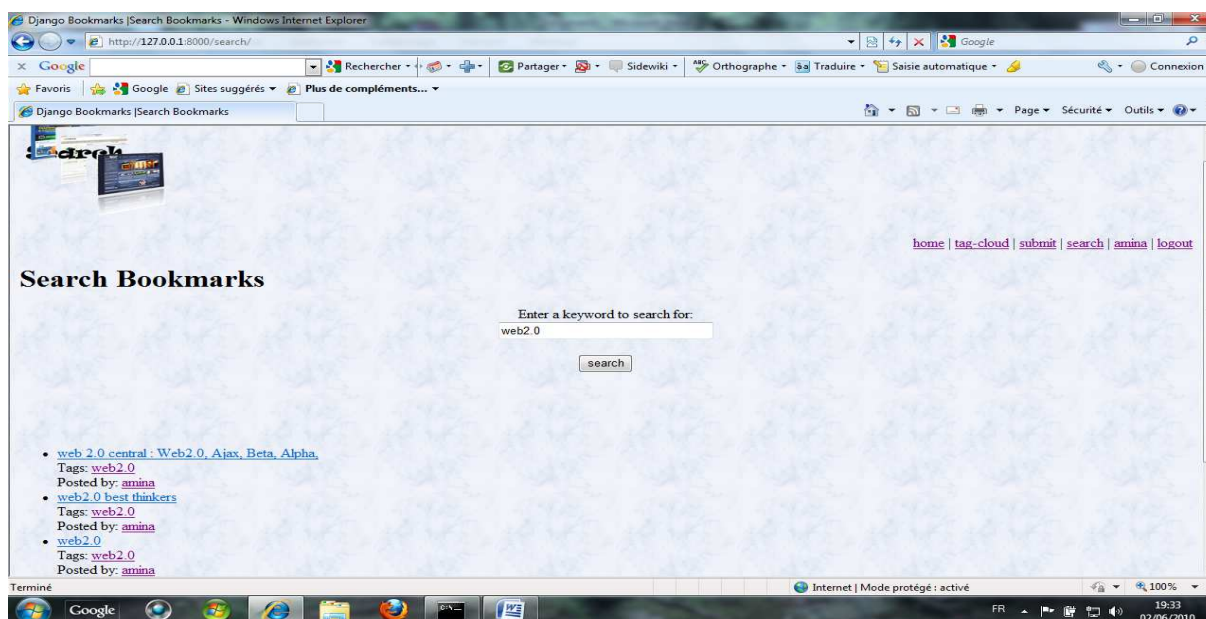


Figure 5.7 : Interface de la recherche

5.3.4. Bookmarks d'utilisateur :

Chaque utilisateur peut avoir tout son propres tags et également il peut les modifier par le clic sur [edit] situer sur la droite de l'URL.

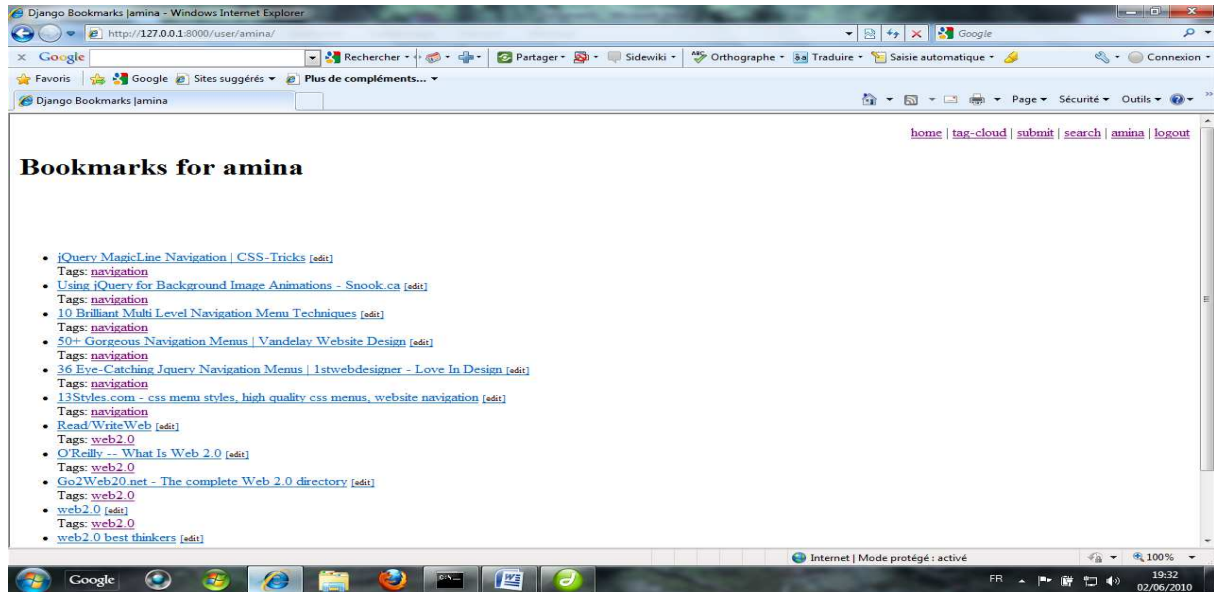


Figure 5.8 : la fenêtre d'affichage des bookmarks d'un utilisateur



Figure5.9 : la fenêtre de la modification d'un Tag

Les sites qui sont sous un Tag peuvent être afficher en cliquant sur le terme correspondant

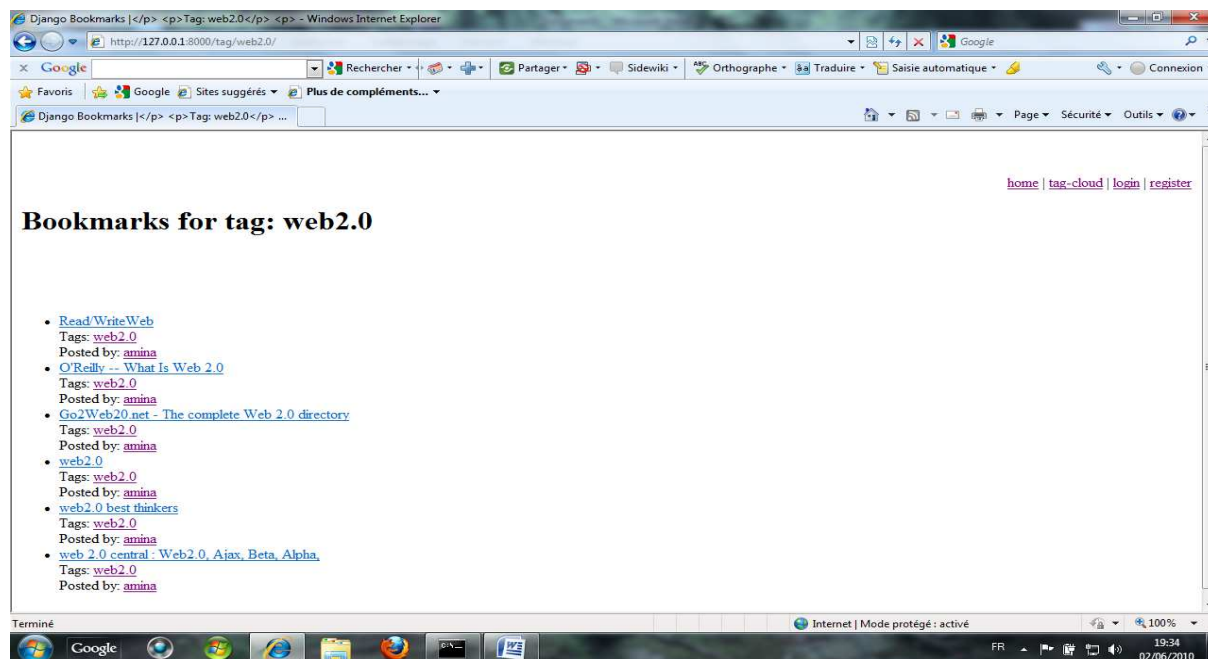


Figure 5.10 : la fenêtre de correspondante au mot web 2.0

Enfin, le système présenter a l'utilisateur un beau nuage de Tag il peut l'utiliser dans la recherche et la navigation sur l'internet.



Figure 5.11 : nuage de Tag de l'application

Lorsque l'utilisateur clique sur une URL ou titre de page dans ce nuage de tags, notre système ouvre le navigateur par défaut du système d'exploitation et affiche cette page.