

# <Document Libre> et le partage du savoir

Jacques Cuvillier  
Association <Document Libre>  
www.documentlibre.org

Que diriez-vous d'avoir dans votre bibliothèque l'ensemble de vos livres et aussi ceux de vos voisins, et même tous ceux de gens dans tout le pays, classés par sujet, numérotés et accessibles en quelques secondes ?

Et non seulement des livres, mais des articles, des images, des vidéos...

On ne rêve pas. C'est ce que permet le numérique à condition de vouloir s'en servir pour mettre au service du plus grand nombre les richesses précieuses dont on ne se défait pas lorsqu'on les donne : la connaissance et le savoir.

Moins d'un an après sa constitution en association, <Document Libre> met en ligne un outil d'indexation et de gestion automatisée de documents accessibles gratuitement sur Internet.

## **Un réseau d'acteurs**

<Document Libre> n'est pas constituée en entité centralisée, mais en un réseau de gestionnaires exerçant chacun leur responsabilité sur une collection de documents. Nul besoin d'un équipement lourd pour tenir ce rôle. Trois petits fichiers html sur son site Internet, et un membre du réseau peut mettre en service la base de données <Document Libre>. L'internaute qui se connecte sur ce site voit apparaître les formulaires interactifs de recherche sous les couleurs et sous le logo de ce site et a le sentiment d'y rester en permanence.

Sauf que... s'il ne trouve pas ce qu'il cherche dans la collection, un clic permet d'élargir sa recherche à toutes les collections du réseau.

## **Un classement thématique universel**

La médiathèque n'est pas destinée qu'à l'enseignement. Elle prétend couvrir les besoins d'un large public. Comme une logique de recherche différente sied à chaque public, sept rayons ont été distingués : activités citoyennes et sociales, arts et loisirs, documentation technologique, enseignement-recherche, illustrations et images, littérature, vulgarisation, vie quotidienne. Chacun de ces rayons est divisé en quelque 12 à 20 thèmes, et au total 110 thèmes différents permettent de couvrir l'ensemble classes que les documentalistes connaissent dans les classifications Dewey ou Cdu [1].

La recherche thématique s'effectue selon trois composantes qui traduisent en fait des éléments standard « Dublin Core » [2] : le sujet (dc.subject) le champ d'application, par exemple, une époque de référence, un lieu (dc.coverage), le public de destination (dc.audience). La conjonction de ces trois éléments crée une classe de documents que l'on nomme « étagère », sur laquelle on peut « instancier » un ou plusieurs objets : des documents appartenant à la même classe thématique.

## **Une sélection par critères spécifiques**

En aval de la recherche thématique, une recherche dite « spécifique » permet de sélectionner les documents par d'autres critères tels que le nom de l'auteur, la date de mise jour, le type, en somme les éléments non thématiques de l'ensemble des éléments « Dublin Core ».

## **Un coup d'oeil, un clic**

C'est la consigne passée aux créateurs du site — au départ des étudiants du département informatique de l'IUT de Nantes — pour concevoir un dispositif de gestion automatisée qui simplifie le plus possible la tâche du gestionnaire. Les fonctionnalités du système sont nombreuses, nous allons en évoquer quelques-unes.

Lorsqu'un internaute propose un nouveau document au gestionnaire d'une collection, il remplit un formulaire en ligne. Suite à ce geste, le gestionnaire de la collection reçoit automatiquement chez lui un courriel avec un hyperlien. En cliquant dessus, il ouvre sur son écran d'ordinateur une page sur laquelle figurent tous les éléments fournis par l'internaute, ainsi que le lien vers le document proposé. Il peut si besoin est intervenir sur le classement proposé par l'auteur, mais si tout est correct, un simple clic va déclencher automatiquement :

- l'immatriculation du document,
- l'enregistrement de ses caractéristiques (et non pas le document lui-même) dans la base de données ,
- la confection d'une fiche descriptive – fichier de métadonnées selon le standard « Dublin Core »,
- l'archivage des données sous forme XML,
- l'envoi d'un courriel de confirmation à l'auteur, avec l'identifiant par lequel son fichier sera désormais nommé et référencé,
- le paramétrage d'un robot de surveillance qui va surveiller que le fichier de l'auteur reste disponible sur Internet, et que ce fichier est toujours le même.
- l'insertion des événements dans des journaux de traitement (logs).

Un bouton « refusé » existe également, dans ce cas l'auteur reçoit un courrier type auquel le webmestre peut adjoindre un post-scriptum.

A partir de ce moment, le document est proposé par le système de recherche. Si le fichier n'est pas accessible sur le site de l'auteur, la mention « indisponible » apparaîtra sur la fiche descriptive du document. Pas question de supporter les « erreur 404 »!

Si l'auteur fait par la suite une mise à jour, le robot détecte la présence d'une nouvelle version et envoie automatiquement un courriel à l'auteur, avec un lien qui lui permet d'ouvrir un formulaire pré-rempli, prêt à recevoir les corrections éventuelles à apporter à la description du document.

## **Pourquoi encombrer son disque dur ?**

Télécharger les fichiers que l'on trouve au fil des recherches sur Internet ne doit plus être un réflexe de collectionneur. Cette pratique, non seulement encombre le disque dur, mais nécessite une méthode de classement efficace. Qui connaît en détail le contenu de ses répertoires ? De plus, on ne profite pas immédiatement des améliorations de la ressource le cas échéant, car personne ne vous avise des mises à jour.

Une autre pratique, autrement plus efficace, est de compter sur une base d'indexation qui permet de retrouver très rapidement un document sur des critères thématiques et de charger la version la plus à jour.

Si l'on veut absolument se créer une bibliothèque locale, il vaut mieux charger son disque avec la description des ressources qu'avec les ressources elles-mêmes. On se contente alors d'enregistrer les fichiers de métadonnées [3] dont c'est précisément le rôle.

## **Quels documents peut-on référencer ?**

La première chose à préciser, c'est que seul l'auteur d'un document (ou son représentant légitime) est à même de référencer une ressource. D'ailleurs, il est impossible de repercuter l'identifiant d'un fichier auquel on n'aurait pas directement accès. Agencé de cette manière, le système fait que

l'auteur d'un document original conserve la maîtrise de son oeuvre. Le principe de l'identifiant qui comporte un champ désignant l'auteur enregistré et l'adoption de la « source unique » de téléchargement sur le site contrôlé par l'auteur en sont une garantie.

Pour le reste, toute la logique de <Document Libre> repose sur l'optimisation de la création de ressources et de l'offre en ligne de documents « libres de droits ». C'est en ce sens que la Charte a été conçue, avec une attitude médiane par rapport aux autres conceptions qui s'opposent. La création ne peut en effet s'accommoder d'une protection totale contre la copie qui interdirait le réemploi et donc freinerait la propagation de la connaissance, mais d'un autre côté, le « pompage intégral » est également de nature à appauvrir la création. Faire un document avec deux moitiés prises ici et là n'est pas faire oeuvre nouvelle !

Bien que la Charte de Document Libre ait été étudiée pour être en adéquation avec les attentes de nombreux auteurs désireux de conserver une grande part de leurs prérogatives, et en adéquation avec le droit français, le système ne l'impose pas. Il impose seulement qu'une cession de droits soit précisée. Dans cette optique, le Charte de Document Libre est à considérer comme une base A MINIMA, c'est à dire que l'auteur peut accorder au lecteur plus de liberté, mais pas moins. Si l'auteur le souhaite, il peut par exemple invoquer la licence « Free Document Licence » ou « Creative Commons ». Ce choix ne devrait pas être fait sur des bases subjectives ou militantes, mais en connaissance de cause [4].

On devrait en particulier avoir à l'esprit que recommander un produit numérique à un « consommateur » s'apparente à un contrat, même si celui-ci s'effectue à titre gratuit. Or :

- l'organisation qui produit l'objet du contrat est soumise à la loi française,
- le document, et en particulier le document pédagogique, n'est pas une ressource comparable à du logiciel.

Le logiciel est censé être transmis entre informaticiens – professionnels ou amateurs – des personnes à même de juger de l'objet du contrat. De plus, s'il est mauvais, l'ordinateur fait la différence : il bogue.

Un document pédagogique par contre, est transmis d'une personne compétente à une personne en situation d'apprentissage, donc vulnérable et à ce titre, protégée par la loi comme consommateur. Par conséquent, les réserves de responsabilité que l'on trouve en particulier dans les licences américaines sont incongrues et risquent fort de n'être pas recevables dans les cas sévères.

Ceci a en particulier la conséquence suivante : chaque fois que la licence adoptée par l'auteur mentionne une clause de non-responsabilité, le webmestre ne peut faire moins que de la répercuter dans la fiche de description du document par la mention « sans garantie », ce qui, à tout prendre, n'est pas de nature à rassurer l'utilisateur...

Jacques Cuvillier

[1]. De fait, les thèmes du rayon « enseignement-recherche » sont assez proches des tables d'entrée de CDU dont certains éléments ont été séparés en deux branches

[2] L'initiative Dublin Core (<http://dublincore.org>) consiste en une recommandation permettant de standardiser les éléments de métadonnées (voir ci-dessous).

[3]. Les métadonnées sont les éléments de texte qui renseignent sur la nature et la teneur d'un document. Chaque document est décrit par un fichier XML de moins de 1Koctet. On peut l'exploiter facilement avec un petit logiciel associé à son navigateur Internet, qui se fondera sur une feuille XSL – sorte de grille de mise en page – pour constituer une sorte de visionneuse de fiches, à utiliser comme des fiches de bibliothèque.

[4] Voir en particulier « Les licences libres et le droit français » de Cyrill Rojinsky et Vincent Grynbaum : [http://www.grynbaum.com/article.php3?id\\_article=10](http://www.grynbaum.com/article.php3?id_article=10)