

Panorama des métadonnées juridiques et de leurs applications en bibliothèque numérique

Lionel Maurel

► **To cite this version:**

Lionel Maurel. Panorama des métadonnées juridiques et de leurs applications en bibliothèque numérique. 2007. sic_00130222

HAL Id: sic_00130222

https://archivesic.ccsd.cnrs.fr/sic_00130222

Preprint submitted on 9 Feb 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PANORAMA DES SYSTEMES DE METADONNEES JURIDIQUES ET DE LEURS APPLICATIONS EN BIBLIOTHEQUE NUMERIQUE

INTRODUCTION

La convergence actuelle d'un faisceau de technologies¹ rend désormais possible l'édification d'importantes bibliothèques virtuelles, voire même peut-être d'envisager la création d'une « bibliothèque virtuelle universelle² », qui regrouperait et diffuserait l'ensemble du savoir humain. Cependant la réalisation de ce rêve babélien risque désormais de se heurter moins à un obstacle technologique qu'à une barrière juridique, qui pourrait s'avérer bien plus redoutable à surmonter. La numérisation et la diffusion des ouvrages met en effet directement en cause les règles du droit d'auteur et notamment, les droits exclusifs de reproduction et de représentation des œuvres qui appartiennent à leurs créateurs ainsi qu'aux différents ayants droit³.

Il existe plusieurs manières de relever le défi des droits des droits d'auteur en matière de numérisation. On sait que depuis plus d'un an Google, le moteur de recherche vedette sur Internet a annoncé le lancement d'un colossal programme de numérisation portant sur quinze millions de livres qui a soulevé beaucoup de polémiques⁴, notamment à cause de doutes quant au respect des règles du droit d'auteur. Cet été, les polémiques opposant Google aux éditeurs ont débouché sur des poursuites judiciaires qui ont fait reculé le géant de l'internet, sans pour autant dissuader plusieurs acteurs économiques – Yahoo!, Microsoft, Amazon – de lancer des projets concurrents. Les bibliothèques numériques existaient déjà avant que Google ne décide d'occuper ce créneau. La première collection numérique, le Projet Gutenberg⁵, remonte à 1971 et les bibliothèques publiques développent depuis longtemps de vastes collections en ligne. En général, les bibliothèques font le choix de se tourner vers des documents patrimoniaux issus du domaine public, afin d'éviter de se heurter à des problèmes de droit d'auteur⁶. Rares sont les bibliothèques qui élargissent leurs collections à des œuvres encore protégées. A titre de contre exemple, on peut citer l'exception notable de la collection numérique de Bibliothèque et Archives nationales du Québec⁷ dont le portail offre l'accès à plusieurs types de documents récents (périodiques,

¹ Ces technologies comprennent les mécanismes de feuilletage automatique des pages, les scanners et les outils de traitement des images, les logiciels de reconnaissance optique des caractères, les langages de structuration des textes, les capacités accrues de stockage des informations, les nouveaux moteurs de recherches, les facilités de téléchargement et de lecture à l'écran des données...etc

² Vladyslav, Frédéric. *Bibliothèque virtuelle universelle : les technologies ne sont plus un obstacle*. Science & Vie, mai 2006, n1064, pp.118-127.

³ Voir à ce sujet : Nguyen Duc Long, Christine. *La numérisation des œuvres : aspects de droit d'auteur et de droits voisins*. Paris : LITEC, 2001. 371 p.

⁴ Pour se faire une idée des débats soulevés par le projet Google Book Search, on peut consulter le dossier de presse constitué par la Bibliothèque nationale de France : BnF. *Le débat autour des projets de numérisation : revue de presse* [en ligne]. Disponible sur :

http://www.bnf.fr/pages/dernmin/com_google.htm.

⁵ Projet Gutenberg : <http://www.gutenberg.net>.

⁶ C'est le cas, par exemple, de Gallica, la collection numérique de la Bibliothèque nationale de France, qui comprend essentiellement des documents libres de droit. La BnF vient cependant d'annoncer qu'elle avait conclu des ententes avec des éditeurs afin d'inclure dans sa collection des titres de presse nationale et régionale encore protégés : <http://gallica.bnf.fr>.

⁷ On peut consulter la collection numérique de BANQ à l'adresse suivante :

http://www.banq.qc.ca/portal/dt/collections/collection_numerique/coll_numerique.jsp

ouvrages de référence, affiches, livres d'artistes...etc) dont les droits ont été libérés par le biais de licences conclues avec les ayants droit.

Si le choix d'incorporer des documents protégés est rare, c'est qu'il implique pour les bibliothèques l'accomplissement de formalités extrêmement lourdes afin de respecter les exigences de la législation sur le droit d'auteur. Identification et recherche des titulaires de droit, négociation de licences, versement de redevances, contrôles éventuels des accès à la collection et des usages sont autant de tâches très coûteuses, en terme financier, mais surtout en termes de temps de travail. A tel point que les obstacles juridiques peuvent s'avérer infranchissables pour les bibliothèques qui ne disposent pas de moyens conséquents.

Heureusement, comme c'est souvent le cas, les problèmes qui naissent lors du passage à l'univers numérique peuvent aussi trouver leur solution grâce à des outils numériques. On a ainsi vu émerger depuis plusieurs années des systèmes permettant de recueillir, de conserver et de gérer, parfois automatiquement, les informations à caractère juridique qui se rattachent aux objets d'une collection numérique. Cet article a pour but d'étudier ces systèmes appelés métadonnées juridiques et de montrer quels bénéfices les bibliothèques peuvent tirer de leur utilisation.

1 METADONNEES JURIDIQUES, DREL ET DRM

Il convient tout d'abord de cerner cette notion de métadonnées juridiques, mais aussi de donner une idée de problématiques plus larges au sein desquelles elles sont impliquées, à savoir les DREL (Digital Rights Expression Languages) et les DRM (Digital Rights Management).

1.1 Les métadonnées juridiques.

Depuis l'antiquité, les bibliothèques ont accomplis des efforts de description et d'indexation des ouvrages visant à faciliter la recherche documentaire. Ces efforts débouchaient sur l'établissement de catalogues de bibliothèques, établis selon les règles de la description bibliographique⁸, qui regroupent des notices décrivant les ouvrages. Avec le passage à l'environnement numérique, l'effort traditionnel de catalogage s'est transformé pour aboutir à la notion de métadonnées. Il s'agit d'informations recueillies à propos de documents primaires, qui servent à décrire, à gérer, à structurer et à conserver objets faisant partie d'une collection numérique. Les métadonnées juridiques constituent une sous-catégorie des métadonnées administratives ou de gestion, spécialement destinées à exprimer les informations juridiques relatives aux documents d'une collection.

Traditionnellement, tant que les supports analogiques ont été prédominants, le monde des bibliothèques a manifesté peu d'intérêt pour ce type d'informations à caractère juridique. Les normes classiques de catalogage (ISBD/MARC) pouvaient inclure la collecte d'informations en rapport avec les questions de propriété intellectuelle comme le nom de l'auteur, la liste des collaborateurs, les noms, lieux, dates d'édition...etc. Mais ces éléments avaient pour but de faciliter la recherche documentaire et non de décrire le statut juridique des œuvres. En réalité, dans la tradition classique du catalogage, il existe même une ambiguïté entre la notion de propriété intellectuelle et deux notions connexes que sont la responsabilité intellectuelle (identification des acteurs impliqués dans la création du contenu d'un document : auteur, co-auteur, préfacier, traducteur, illustrateur...) et l'adresse bibliographique (identification de l'entité responsable de la diffusion de l'œuvre, ainsi que des dates et lieux d'édition).

De ces ambiguïtés peuvent découler des confusions. Les informations contenues dans les notices bibliographiques ne permettent pas réellement de déterminer avec précision l'identité des titulaires de droit. L'auteur ou l'éditeur d'une œuvre ne sont pas forcément détenteurs des droits sur l'œuvre, surtout

⁸ Ces normes de catalogage ont été fixées au niveau international par des standards ISBD (International Standard Bibliographic Description). Lorsque les catalogues de bibliothèques ont été informatisés, les informations bibliographiques ont été transcrites dans un format lisible par les machines : MARC (Machine Readable Cataloguing), qui se décline en différentes versions nationales.

dans un système comme celui du droit canadien dans lequel les cas de dissociation entre l'auteur et le titulaire des droits sont relativement fréquents⁹.

On peut à la rigueur estimer que les informations décrivant la responsabilité intellectuelle peuvent permettre de pister les titulaires du droit moral sur une oeuvre, mais ce n'est pas toujours suffisant (le titulaire a pu renoncer par contrat à exercer son droit moral ou ses droits ont pu s'éteindre lorsque l'oeuvre est tombée dans le domaine public).

Il s'ensuit que les informations juridiques disponibles dans les catalogues traditionnels sont relativement minces. On verra plus loin que ces lacunes sont particulièrement gênantes, car elles contraignent les bibliothèques à repartir de zéro pour établir des fichiers de données juridiques à partir de métadonnées juridiques.

L'indigence des données juridiques dans les formes traditionnelles de description bibliographique a conduit la communauté des bibliothécaires et des professionnels de l'information à développer des systèmes de métadonnées qui prennent mieux en compte la dimension de la propriété intellectuelle. C'est le cas depuis 1995 du Dublin Core, un des premiers systèmes de métadonnées, dont l'une des grandes originalités est d'inclure à côté des éléments de Contenu et d'Identification des éléments de Propriété intellectuelle¹⁰.

Cependant, c'est surtout avec le développement d'Internet et du commerce électronique que le besoin de développer des systèmes de données juridiques s'est fait le plus fortement sentir. Les innovations les plus importantes dans ce domaine proviennent du secteur commercial. C'est ce qui a conduit à la création des DREL.

1.2 Les DREL (Digital Rights Expression Languages)¹¹

C'est la nécessité de créer des environnements sécurisés pour permettre le développement du commerce électronique qui a conduit au développement des langages d'expression des droits numériques. L'idée est de pouvoir décrire précisément le statut juridique des oeuvres numériques (tâche souvent complexe), d'exprimer les dispositions découlant d'accords contractuels et de licences de manière à pouvoir mettre en oeuvre de manière efficace des systèmes de contrôles des accès et des usages.

La conception des DREL a débuté dans les années 90, avec les travaux effectués par Mark Stefik pour le compte de la firme Xerox. En 1998, Microsoft et Xerox se sont associés dans le cadre de la société ContentGuard pour créer le XrML (eXtensible rights Markup Language) appelé à devenir un standard important dans le domaine. C'est en utilisant cette base qu'a été créé en 2003 le MPEG-21/5, standard qui tend peu à peu à s'imposer et qui a même reçu une certification ISO en 2004¹². Parallèlement, une multitude de standards proliféraient pour s'adapter aux besoins de secteurs particuliers, sans atteindre l'importance du MPEG-21.

L'autre standard important en matière de DREL est le ODRL (Open Digital Rights Language) qui a été développé en Australie depuis 2000 sous la forme d'un standard ouvert. Il présente la particularité de pouvoir être adopté aussi bien par le secteur commercial que par des bibliothèques en raison de sa souplesse.

⁹ C'est le cas par exemples des oeuvres créées par des employés, des photographies, gravures et portraits réalisés sur commande...etc

¹⁰ Voir : Dublin Core Metadata Initiative. *DCMI Metadata Terms* [**en ligne**]. Disponible sur : <http://dublincore.org/documents/dcmi-terms/>

¹¹ Les études portant sur les métadonnées juridiques en bibliothèque et sur les DREL sont assez rares. Nous nous sommes beaucoup appuyé sur les travaux de Karen Coyle, ancienne bibliothécaire à l'Université de Californie et auteure d'un important rapport sur les DREL pour la bibliothèque du Congrès. Coyle, Karen. *Rights expression languages. A report for the library of Congress. Février 2004* [**en ligne**]. Disponible sur : http://www.loc.gov/standards/Coylereport_final1single.pdf.

¹² ISO/IEC FDIS 21000-5 :2003(E).

Le monde des bibliothèques a finalement lui aussi mis au point des normes en matière de DREL, car les standards développés par le secteur commercial correspondent assez mal à ses besoins. Le Dublin Core a ainsi été enrichi par le biais de nouveaux éléments et qualificatifs pour mieux prendre en compte les questions de propriété intellectuelle. Un volet juridique a été donné au standard de métadonnées METS développé par la Bibliothèque du Congrès (METSRights). Enfin, le domaine de la Culture Libre a lui aussi mis au point depuis 2002 un système original d'expression des droits, qui n'est pas sans susciter de l'intérêt dans le monde des bibliothèques : les Creative Commons.

Il faut préciser qu'à l'heure actuelle, ces systèmes sont encore pour la plupart à un stade expérimental de développement. On compte peu d'implémentations réelles, et moins encore dans le monde des bibliothèques. Mais les bibliothèques n'ont pour l'instant pas exprimé de manière nette les besoins qu'elles pouvaient avoir en la matière.

On ne peut comprendre l'enjeu réel des langages d'expression des droits sans évoquer la question des DRM.

1.3 Les DRM (Digital Rights Management) :

Les DREL (du moins ceux produits par la sphère commerciale) n'ont pas seulement pour but de collecter des informations juridiques, mais aussi de se servir de ces données pour mettre en place des environnements sécurisés qui préviennent les utilisations illégales des œuvres, en servant de support à des systèmes de contrôle efficaces des accès et des usages. C'est ce que l'on nomme les DRM, traduit par systèmes de gestion numérique des droits¹³.

Les DREL ont ainsi vocation à être exprimés dans des langages lisibles par les machines (XML notamment¹⁴), de manière à pouvoir être mis automatiquement en application. Tous les DREL ne permettent cependant pas de tels emplois.

Les DRM soulèvent actuellement beaucoup de polémiques¹⁵ et ont généralement mauvaise presse, particulièrement dans le monde des bibliothèques. On considère en effet qu'il pourrait servir à mettre en place des systèmes coercitifs qui nuiraient grandement à la circulation des œuvres numériques, voire porteraient atteinte à la vie privée et à l'exercice de certaines libertés. Les bibliothécaires craignent également que le développement de ce genre de systèmes ne nuise à terme à l'accomplissement de leurs missions¹⁶.

¹³ Il existe tout un débat pour savoir si l'on doit dire Gestion numérique des droits ou Gestion des droits numériques. Il nous semble qu'il vaut mieux parler de gestion numérique des droits, car les droits d'auteur ne changent pas de nature quand on passe dans l'univers numérique. C'est bien leur gestion qui utilise des technologies numériques.

¹⁴ XML pour eXtensible Markup Language. Il s'agit d'un langage de programmation universel qui structure les documents et les données pour en faciliter la transmission et l'extraction des informations. Ce langage joue un rôle grandissant sur Internet en général et dans la construction des bibliothèques virtuelles en particulier.

¹⁵ Ce fut notamment un des grands sujets débattus avec virulence lors de l'adoption de la Loi DADVSI en France cet été. Voir la référence suivante pour un exemple assez représentatif de l'état d'esprit en France sur la question des DRM : Barthe, Emmanuel. *DRM et documents : les risques d'un futur proche*. 26/04/2005 [en ligne]. Disponible sur : http://www.servicedoc.info/article.php3?id_article=174. Le projet de loi C-60 sur la réforme du droit d'auteur se proposait de donner une existence légale aux DRM. Après l'abandon de ce projet de réforme, le Canada n'a toujours pas statué sur leur sort. Les DRM soulèvent aussi au Canada des inquiétudes : Cf. Morin, Phillippe. *Les mesures techniques de protection du droit d'auteur – Aperçus des conséquences possibles en droit canadien : copie pour usage privé et exceptions au droit d'auteur*. Les cahiers de Propriété Intellectuelle, Vol. 17, n 2, mai 2005, pp. 277-337.

¹⁶ Dans ce sens : Coyle, Karen. *Digital Rights Management* [en ligne]. Disponible sur : http://www.kcoyle.net/drm_basics3.html.

Il faut cependant être nuancé et distinguer les systèmes de gestion numérique des droits (DRM) des mesures techniques de protection (MTP). Ces mesures représentent seulement un des aspects des systèmes de gestion des droits. Il s'agit de mesures de cryptage ou de verrouillage directement implantées dans les œuvres (CD, DVD, Logiciels) qui peuvent être particulièrement gênantes pour les bibliothèques. Mais les DRM ne se limitent pas à ces mesures contraignantes, et, pris dans toutes leurs dimensions, ils peuvent aider à gérer de manière plus fluide les questions de propriété intellectuelle, qui représentent une charge de plus en plus écrasante pour les bibliothèques.

On ne s'interdira donc pas dans ce rapport de chercher quels bénéfices les bibliothèques peuvent tirer des DRM.

2. METADONNEES JURIDIQUES ET BESOINS DES BIBLIOTHEQUES NUMERIQUES

2.1 Quels sont les besoins d'une bibliothèque numérique en la matière?

Nous l'avons vu, les règles issues de la loi sur le droit d'auteur imposent des formalités extrêmement lourdes dont la charge peut devenir un obstacle non négligeable au développement des collections numériques. Cet obstacle peut même devenir infranchissable si l'on en reste à une gestion papier des questions de droit d'auteur. Les métadonnées juridiques permettent d'envisager une gestion beaucoup plus fiable et efficace, voire même un traitement automatisé de l'information juridique.

Il importe avant d'étudier certains des systèmes actuellement disponibles de recenser les besoins qui peuvent être ceux d'une bibliothèque numérique en matière de gestion des droits :

- Exprimer de manière claire le statut juridique des œuvres (œuvre protégée, domaine public...), ainsi que la titularité des droits (quels sont le ou les titulaires de droits sur une œuvre ?)
- Recenser et conserver des informations utiles dans le cadre de la recherche d'ayants droit.
- Recueillir les informations relatives à des licences conclues avec des titulaires de droit, les conserver et en faciliter l'accès, tout en protégeant les données confidentielles.
- Gérer des informations relatives aux redevances versées pour diffuser les œuvres.
- Délivrer des informations juridiques claires et fiables aux utilisateurs de la collection numérique concernant leurs droits et obligations.
- Permettre l'échange et le partage des informations juridiques dans le cadre d'un travail en réseau.
- Contrôler les accès à la collection numérique.
- Contrôler les usages et les réutilisations de la collection numérique.

Les métadonnées juridiques, les DREL et les DRM permettent à des degrés divers de satisfaire ces besoins.

2.2 Objectif de cette étude

Le but de cette étude est d'abord de décrire succinctement les systèmes de métadonnées les plus importants actuellement. On envisagera ainsi successivement cinq d'entre eux :

- Le volet Propriété Intellectuelle du Dublin Core.
- METSRights.
- Les CreativeCommons.
- ODRL et MPEG-21/5

Bien d'autres systèmes existent, la matière est même particulièrement foisonnante¹⁷, mais cet échantillon présente à la fois l'intérêt d'être suffisamment représentatif des différentes tendances et de correspondre plus directement aux besoins d'une bibliothèque numérique.

Pour chacun de ces systèmes, on s'efforcera, après l'avoir décrit, de chercher à savoir de quelle manière il peut satisfaire les besoins d'une bibliothèque numérique.

PREMIERE PARTIE : LE VOLET PROPRIETE INTELLECTUELLE DU DUBLIN CORE

1 DESCRIPTION

A l'origine en 1995, le Dublin Core¹⁸ proposait parmi ces quinze éléments quatre champs de données relatifs à la propriété intellectuelle ;

- **Creator/Créateur** : l'entité principale responsable de la création du contenu de la ressource.
- **Contributor/Collaborateur** : Une entité qui a contribué à la création du contenu de la ressource.
- **Publisher/Editeur** : l'entité qui rend la ressource disponible.
- **Rights/Gestion des droits** : informations sur les droits et au sujet de la ressource.

Ce bloc « Propriété intellectuelle » a été salué à l'époque comme une entreprise louable et comme un enrichissement par rapport à la tradition du catalogage classique. Cependant, il n'est pas exempt de défauts :

- La question de la propriété intellectuelle est mal dégagée de la responsabilité intellectuelle. Rien ne garantit que le créateur, les contributeurs ou l'éditeur seront les titulaires des droits sur la ressource. Les droits ont pu appartenir dès l'origine à une autre personne ou avoir été cédés par la suite de manière contractuelle. On peut à la rigueur pister les titulaires de droit moral avec ces éléments (et encore), pas ceux qui possèdent les droits patrimoniaux.

- On peut penser que l'élément Rights était destiné à faire figurer l'identité du titulaire de droit, mais il faut avouer que la signification de ce champ n'est pas très claire (commentaire du Dublin Core Metadata Initiative – DCMI- : « typiquement, un éléments Droits contiendra un état du droit à gérer cette ressource, ou la référence à un service fournissant cette information).

Ces lacunes ont conduit les gestionnaires du Dublin Core à développer de nouveaux éléments et de nouveaux qualificatifs de raffinement qui présentent plus d'un intérêt.

Dans la catégorie des nouveaux éléments :

- **Rights Holder/Titulaire des droits** : une personne ou une organisation qui possède ou qui gère les droits sur la ressource.
- **Provenance/Possesseur** : une indication des changements intervenus dans la propriété ou dans la garde de la ressource depuis sa création.

Dans la catégorie des qualificatifs de raffinement (attachés à l'élément Rights) :

¹⁷ On peut citer par exemple ERMI (Electronic Ressources Management Initiative), FDRM (Federated Digital Resource Management), PRISM (Publishing Requierements for Industry Standard Metadata ou XACML (eXtended Acces Control Markup Language). On trouvera des informations sur ces systèmes à l'adresse suivante : <http://xml.coverpages.org>

¹⁸ Une description rapide du Dublin Core et de ses éléments : Bibliothèque de l'Université Laval. *Guide d'utilisation du Dublin Core. Les éléments de base : propriété intellectuelle* [en ligne]. Disponible sur : <http://www.bibl.ulaval.ca/DublinCore/usageguide-sectc-20000716fr.htm>

- **AccesRights/Droits d'accès** : information indiquant qui peut accéder à la ressource ou son niveau de sécurité.
- **Licence** : un document à valeur légale qui autorise officiellement certains usages de la ressource

On voit immédiatement l'intérêt d'un tel apport. Avec l'élément Rights Holder, les questions de propriété intellectuelle sont traitées de manière distincte de celles de responsabilité intellectuelle. De plus, la présence de ce nouvel élément décharge l'élément Rights/Gestion des droits qui est libre pour de nouveaux usages (accès/licence).

Par ailleurs, d'autres éléments qui ne figurent pas dans le bloc Propriété Intellectuelle, mais dans les catégories Contenu ou Identification, peuvent néanmoins avoir une utilité dans le cadre de la gestion des droits :

- **Date** : une date associée à un événement du cycle de vie de la ressource. Peut se raffiner en Available, Created, Date Copyrighted, Issued, Modified. Ces informations peuvent servir à calculer la durée de protection des œuvres, qui prend appui selon les cas sur des dates différentes.
- **Title** : un nom donné à la ressource. Utile pour l'identification des œuvres.
- **Type** : la nature ou le genre du contenu de la ressource. Information utile car les règles applicables en matière de droit d'auteur sont susceptibles de varier en fonction de la nature du document (imprimé, enregistrements sonores, photographies, gravures, films, documents nés numériques...etc)

2 Comparaison avec les besoins d'une bibliothèque numérique.

Dans le milieu des bibliothèques, on déplore souvent le côté fruste et réducteur du Dublin Core, mais force est de constater qu'en matière de propriété intellectuelle, ce langage a atteint un niveau de développement qui peut le rendre intéressant.

2.1 Besoins en matière de description du statut juridique des ressources et de la titularité des droits.

La combinaison des champs Creator/Contributor/Publisher/Rightsholder permet une description relativement fine de la titularité des droits sur les ressources, en dissociant bien la responsabilité intellectuelle de la propriété intellectuelle. Ces éléments multiples de description peuvent être utiles dans les cas (fréquents) où plusieurs agents sont titulaires des droits sur une même œuvre (affiches, livres d'artiste...). On peut aussi prévoir une valeur spéciale du type « Domaine public » que l'on pourra attribuer à l'élément RightsHolder pour les œuvres qui sont devenues libres de droit avec le temps.

On peut imaginer exploiter le Dublin Core en liaison avec un système de gestion intégrée des documents¹⁹(GID). Un fichier général des titulaires de droits peut ainsi être mis en place en attribuant un numéro aux titulaires de droit. Ce numéro autorise à faire un lien entre les métadonnées en Dublin Core qui recueilleront seulement le nom du titulaire et le fichier qui pourra inclure de plus amples informations (notamment des données nominatives pour permettre de rejoindre les personnes, pouvant par ce biais être gardées confidentielles).

Le DCMI préconise l'emploi d'un URI (Uniform Resource Identifier) pour désigner les titulaires de droit.

¹⁹ La gestion intégrée des documents (GID) correspond au passage à une gestion informatisée des documents d'archive internes produits par une organisation. Cf. Roberge, Michel. *Gestion intégrée des documents (GID)* [en ligne] . Disponible sur : <http://gestionintgre.blogspot.com>

2.2 Besoins en matière de gestion des licences et des redevances.

L'élément Licence permet d'indiquer si un accord a été conclu avec les ayants droit concernant l'usage de la ressource. On imagine qu'un codage spécial permettrait d'indiquer expressément quels usages ont été autorisés par la licence. On peut aussi prévoir un codage qui indique pour combien de temps la ressource est valable et si elle implique que soient versées des redevances, selon telles ou telles modalités (montant, sommes forfaitaires, redevances périodiques).

Ici encore, le DCMI préconise d'attribuer aux licences un numéro d'identification de type URI. Cette démarche peut là aussi être utile dans le cadre d'un système de GID. Ce numéro permettrait de faire le lien avec une base de données comportant les licences numérisées et/ou des informations synthétiques sur ces licences. Un renvoi peut également être prévu vers le fichier des titulaires de droit, permettant des ponts pratiques entre ces informations.

2.3 Besoins en matière de gestion des accès et des usages

Le qualificatif AccesRights permet de faire figurer des limitations d'accès particulières, notamment entre les ressources qui peuvent être accessibles en ligne et celles dont l'accès est réservé sur place. Un codage spécial peut faire le lien avec les licences lorsque ces contraintes trouvent leur origine dans un contrat.

Les questions d'accès et d'usage peuvent commencer à recevoir un traitement grâce aux éléments AccesRights et Licence, mais il faut bien reconnaître que les possibilités sont encore assez limitées.

Au final, le Dublin Core couvre assez bien les besoins d'une bibliothèque numérique, tant que ceux-ci ne sont pas trop complexes. On peut bien sûr profiter de la souplesse de ce langage pour l'enrichir en interne, mais on perd alors en interopérabilité, qui reste par ailleurs un des intérêts majeurs du Dublin Core, dans le cadre du protocole OAI-PMH²⁰.

DEUXIEME PARTIE : METSRights

1. DESCRIPTION

METS²¹ (Metadata Encoding and Transmission Initiative) est un standard de métadonnées, développé aux Etats-Unis par la Digital Library Federation. Sa mise à jour est assurée par la Bibliothèque du Congrès. Son but est de fournir un cadre permettant d'exprimer en XML des informations descriptives, administratives et structurelles à propos de ressources numériques. Il est appuyé par une communauté regroupant plutôt des institutions universitaires et de grandes bibliothèques, ce qui lui donne une certaine orientation. Il a plutôt été conçu pour des établissements publiant des archives ouvertes ou diffusant des documents à des fins éducatives. Mais à titre de contre-exemple, c'est aussi le système de métadonnées qui sert actuellement en France dans le cadre de la gestion de Gallica, la collection numérique de la BnF. METS est compatible avec le protocole OAI-PMH.

²⁰ Ce protocole d'échange permet la collecte et l'échange des métadonnées sur Internet, ce qui permet de faciliter la recherche documentaire. Cf. Open Archives Initiative. *The Open Archives Initiative Protocol for metadata harvesting* [**en ligne**]. Disponible sur :

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

²¹ Library of Congress. *METS : Metadata Encoding and Transmission Standard* [**en ligne**]. Disponible sur : <http://www.loc.gov/standards/mets/>

METS a récemment reçu une extension spécialement dédiée à la collecte d'informations à caractère juridique, intitulé : METSRights²². L'idée est de permettre d'adjoindre aux métadonnées METS attachée à une ressource une « RightsDeclaration » regroupant un minimum de données administratives à propos des droits de propriété intellectuelle. On va voir que ce « minimum » est tout de même relativement développé. METSR a été conçu pour collecter des informations juridiques, mais pas pour les faire appliquer automatiquement par un système de sécurité de type DRM. Ce sont plutôt des informations destinées à être réutilisées par des humains.

METSR fonctionne à partir d'« ElementTypes », agencés dans un ordre fixe et dont le sens peut être précisé par des éléments de raffinement.

On trouve d'abord un élément RightsDeclaration, destiné à regrouper sommairement des informations sur le statut juridique de la ressource, pouvant être affichées sous forme de notice à l'intention des utilisateurs.

Une série d'éléments développent par la suite ces informations de manière détaillée :

- **RightsHolder** : On indique ici la personne ou l'organisation qui porte les droits d'auteur pour la ressource considérée, sachant que METSR fait une intéressante distinction entre le RightsHolder et le CopyrightHolder. Le premier est le titulaire des droits, tandis que le second est celui qui gère les droits, comme dans le cas d'une société collective de gestion qui va représenter un auteur. Cet élément peut être répété lorsqu'il y a plusieurs titulaires de droit sur une œuvre. Le RightsHolder peut être identifié à l'aide d'un identifiant, le RightsHolderID.
On précise le sens de ce champ de la façon suivante :
 - *RightsHolderName* : le nom de la personne ou de l'organisation qui porte les droits, lié au RightsHolderID.
 - *RightsHolderComment* : une zone de commentaires libres qui peut être utilisée par exemple pour préciser la fonction du RightsHolder vis-à-vis de l'œuvre (auteur, éditeur, traducteur...etc)
 - *RightsHolderContact* : on indique dans ce champ la façon dont on peut joindre en cas de besoin le titulaire des droits. C'est ici que l'on peut indiquer si le titulaire des droits est représenté par une autre personne. Tout un ensemble de sous-champs sont dédiés aux informations permettant de faciliter le contact : Designation, Address, Phone (qui peut être Fax, Mobile ou Business) et Email.
- **RightsCategory** : On mentionne ici si l'œuvre est Copyrighted (protégée), Licensed (fait l'objet d'une licence), Public Domain (tombée dans le domaine public), Contractual (comme dans le cas d'une ressource achetée type périodique électronique) ou Other. Le sens de Other peut être précisé dans un champ OtherCategoryType.
- **Context** : ce champ sert à décrire qui peut faire quoi avec la ressource et selon quelles conditions. On peut donner un identifiant à chaque définition de Context et le lier à un RightsHolderID. Ce champ se subdivise en :
 - *ContextClass* : un groupe d'utilisateur qui peut être : Academic User, General Public, Repository Manager, Managed GRP, Institutional Affiliate ou Other.
 - *UserName* et *UserType* : permettent de désigner nommément un groupe ou une personne pour leur réserver l'accès à une ressource.
 - *Permissions* : indique quels types d'usage sont autorisés parmi la liste suivante : DISCOVER, DISPLAY, COPY, DUPLICATE, MODIFY, DELETE, PRINT, OTHER.
 - *Constraints* : permet de soumettre l'usage à un certain nombre de conditions : QUALITY, FORMAT, UNIT, WATERMARK, PAYMENT, COUNT, ATTRIBUTION, RE-USE, TIME, TRANSFERPERMISSIONS et OTHER.

2. Comparaison avec les besoins d'une bibliothèque numérique.

²² Library of Congress. *Index of standards/Rights* [en ligne]. Disponible sur : <http://www.loc.gov/standards/rights/METSRights.xsd>.

2.1 Besoins en matière de description du statut juridique de la ressource et de la titularité des droits

L'élément RightsCategory satisfait *a minima* la fonction de description du statut juridique de la ressource (protégée, dans le domaine public...). En revanche, l'élément RightsHolder permet de décrire de manière détaillée le titulaire des droits. Les champs qui permettent de rassembler des données pour recontacter les titulaires des droits ou leurs agents peuvent être très utiles. On peut aussi imaginer que même lorsqu'on ne parvient pas à retrouver le titulaire de droit après des recherches, on puisse tout de même garder une trace des recherches en les incluant dans le champ RightsHolderContact, très utile par exemple pour constituer des dossiers de demande de licence pour titulaire de droit introuvable auprès de la Commission fédérale du droit d'auteur²³.

METSR paraît particulièrement bien adapté de ce point de vue pour former un fichier de métadonnées extérieures aux ressources numériques de la collection, qui pourrait servir de base à la création d'un fichier des titulaires de droits. La possibilité d'attribuer un identifiant au titulaire permet d'envisager des applications en matière de GID.

Par contre, ces informations sont des données nominatives à caractère personnel qui doivent faire l'objet d'un accès contrôlé en vertu de la loi. On ne peut les mettre à la disposition de quiconque. Il faudrait donc prévoir une façon de réserver l'accès au champ RightsHolderContact.

2.2 Besoins en matière de gestion des licences et des redevances

METSR est un peu plus décevant de ce point de vue.

On peut indiquer avec l'élément RightsCategory si une licence a été conclue avec un ayant droit à propos de la ressource (Licensed), mais on ne peut préciser la teneur de la licence, ni ses caractéristiques, ni même lui attribuer un numéro.

La gestion des redevances n'est tout simplement pas prévue par METSR.

2.3 Besoins en matière de gestion des accès et des usages.

METSR est bien développé sur ces points, mais il révèle son origine « universitaire ».

La gestion des accès est bien couverte par les champs ContextClass, UserName et UserType, ce qui peut s'avérer très utile dans le contexte d'une bibliothèque universitaire par exemple qui offre un accès différencié aux ressources en fonction du statut de l'utilisateur (professeurs, bibliothécaires, étudiants de différentes années, utilisateurs extérieurs...etc).

les éléments de gestion des usages sont aussi présents grâce à la combinaison des champs Permissions et Constraints. On pourrait imaginer par ce biais que certaines ressources pourraient seulement être consultées et non copiées, dupliquées ou imprimées. Il est même possible d'envisager grâce à l'élément Constraints des limites plus complexes, comme « copier seulement X pages d'un document », « pouvoir consulter

²³ L'article 77 de la Loi sur le Droit d'Auteur prévoit que la Commission du Droit d'Auteur peut délivrer sur demande une licence d'utilisation lorsque les titulaires de droits d'une œuvre ne peuvent être identifiés ou contactés. La loi exige que le demandeur fasse la preuve qu'il a accompli des efforts raisonnables pour retrouver le titulaire de droits. On comprend dès lors tout l'intérêt de pouvoir garder une trace des recherches effectuées, d'autant plus que les cas où les titulaires de droits ne peuvent être retrouvés sont relativement fréquents. Cf. Commission fédérale du droit d'auteur. *Titulaires de droit introuvables* : Brochure [en ligne]. Disponible sur : <http://www.cb-cda.gc.ca/unlocatable/brochure-f.html>.

seulement un document pendant X unités de temps, puis le copier », « arrêter de pouvoir consulter un document après X unités de temps », « copier seulement avec tel ou tel format ou telle ou telle qualité », « consulter ou copier après paiement ou à partir d'un compte d'utilisateur »...etc

Certains éléments de Constraints peuvent aussi servir de support pour indiquer des « bonnes pratiques » de ré-utilisation des œuvres de la collection numérique d'une bibliothèque. Ex : ATTRIBUTION = réutiliser à condition d'indiquer le nom de l'auteur ou d'indiquer que l'œuvre provient de la collection de la bibliothèque, RE-USE = empêcher la réutilisation de certaines œuvres (celles encore protégées des droits d'auteur), WATERMARK = inclure un « tatouage » dans les œuvres de manière à les marquer ou ajouter un filigrane du type « Issu de la collection numérique de la bibliothèque X » qui s'affichera à chaque impression.

La possibilité d'afficher une notice propre à chaque œuvre qui informe de manière rapide les usagers sur leurs droits paraît aussi très intéressante, dans l'optique de développer une « pédagogie » des droits d'auteur. Les bibliothèques numériques délivrent en général des informations à leurs usagers concernant leurs droits et obligations, par le biais d'une page de leur portail. Ces avertissements sur les droits d'auteur peuvent parfois être peu visible et ils ne tiennent pas compte des différences de statut entre les œuvres (protégée, dans le domaine public...etc).

Un outil comme METSR peut permettre une gestion plus étroite des usages, ou du moins d'orienter les usages dans telles ou telles directions en fournissant des informations *ad hoc*, pour chaque ressource, aux utilisateurs.

Il faut toutefois se souvenir que METSR n'est a priori pas conçu pour mettre en œuvre directement des systèmes automatisés de contrôle des usages (DRM). Il s'agit plus d'un vecteur d'informations à caractère juridique à usage interne ou externe que d'un système de verrouillage.

TROISIEME PARTIE : LES CREATIVE COMMONS

1. DESCRIPTION

Les Creative Commons ont été créées en 2001 par le juriste Lawrence Lessig l'un des principaux animateurs du mouvement de la Free Culture, qui revendique le passage à une nouvelle conception du droit d'auteur, mieux adaptée aux nouvelles possibilités d'échange et de création offertes par l'environnement numérique. Les Creative Commons sont depuis développés par une communauté²⁴ qui possède des prolongements dans de nombreux pays afin de les adapter à chaque système juridique. un groupe d'avocats et d'informaticiens américains de la Stanford Law School, revendiquant leur allégeance au mouvement de la Culture Libre. Parmi ce groupe, on trouve en particulier le juriste Lawrence Lessing qui est un des principaux animateurs de la Free Culture. L'idée de base qui sous-tend les Creative Commons²⁵ est de créer un système simple d'expression des droits qui permette une large diffusion des œuvres via Internet. Les Creative Commons peuvent être assimilés à un système de licences a priori que les auteurs associent à leurs œuvres et qui indiquent à l'avance quels types d'usages ils autorisent ou non. Ce système inverse quelque peu la logique habituelle du fonctionnement des droits d'auteur. Dans l'approche traditionnelle, les œuvres bénéficient dès leur création d'une protection (Copyright) qui nécessite que l'on demande autorisation à leur auteur pour les utiliser. Si l'œuvre est diffusée sur Internet, c'est seulement, selon

²⁴ Voir le site officiel des Creative Commons : Creative Commons [**en ligne**]. Disponible sur : <http://creativecommons.org/> , ainsi que le site canadien : Creative Commons Canada [**en ligne**]. Disponible sur : <http://creativecommons.ca/>.

²⁵ Pour une présentation rapide des Creative Commons réalisée par la bibliothèque de l'UQAM : Bibliothèques UQAM. *Dossier - Creative Commons ou « certains droits réservés »* [**en ligne**]. Disponible sur : http://www.bibliotheques.uqam.ca/informations/bibliocliq/dossiers/creative_commons.html

l'expression consacrée : « tous droits réservés », symbolisé par le signe bien connu ©. Avec les Creative Commons, l'auteur décide à l'avance de se réserver seulement certains droits et d'abandonner les autres, de façon à permettre une circulation et une réutilisation plus aisée de son œuvre.

Les Creative Commons constituent bel et bien un DREL (Digital Rights Expression Language), sauf qu'ils fonctionnent de manière inverse des autres langages étudiés dans cette étude : au lieu d'être prohibitifs, d'interdire des usages, ils sont permissifs. Dans le système classique du Copyright, la règle est : « tout ce qui n'est pas permis est interdit ». Avec les Creative Commons, la règle devient : « tout ce qui n'est pas interdit est permis ».

Il faut ajouter que toutes les interdictions prévues par les Creative Commons peuvent être levées, à condition d'obtenir l'autorisation préalable du titulaire des droits.

Les Creative Commons peuvent paraître assez éloignés des préoccupations d'une bibliothèque numérique publique, sauf que des institutions très importantes recommandent leur utilisation, souvent couplée avec l'emploi du Dublin Core. C'est le cas notamment du très important projet Michael de répertoire des ressources numériques européennes²⁶, ou encore du logiciel DSpace²⁷, très utilisé dans le monde universitaire comme outil de diffusion des collections numériques.

En fait, comme les logiciels libres en leur temps ont pu trouver emploi dans les bibliothèques publiques, les Creative Commons, issus de la Culture Libre, peuvent être employés avec bénéfice en bibliothèque. Creative Commons a d'ailleurs reçu des développements pour s'adapter au domaine des archives (Creative Archive²⁸) et au domaine de la diffusion des archives ouvertes en matière scientifique (Science Commons²⁹).

A la base du système on trouve un système d'icônes qui permettent par combinaison aux titulaires des droits d'auteur de construire des licences :

	Paternité : Autorise la redistribution libre de l'œuvre, en autant que le créateur original est cité.
	Pas d'Utilisation Commerciale : L'œuvre peut être reproduite, distribuée, communiquée ou modifiée mais elle ne peut être réutilisée à des fins commerciales.
	Pas de Modification : L'œuvre peut être reproduite, distribuée, communiquée ou utilisée à des fins commerciales mais elle ne peut être modifiée, transformée ou adaptée.
	Partage des Conditions Initiales à l'Identique : L'œuvre peut être modifiée, transformée ou adaptée mais la création qui en résulte doit être distribuée sous un contrat identique à celui de l'œuvre originale.

En combinant selon son choix ces symboles, l'auteur peut former 11 types de licences différentes dont 6 sont recommandées par la Creative Commons Initiative (ici présentées de la plus restrictive à la plus permissive).

²⁶ MICHAEL. *Multilingual Inventory of Cultural Heritage in Europe* [en ligne]. Disponible sur : http://www.michael-culture.org/index_f.html

²⁷ DSpace Federation [en ligne]. Disponible sur : <http://www.dspace.org/>

²⁸ Creative Archive [en ligne]. Disponible sur : <http://creativearchive.bbc.co.uk/>.

²⁹ Sciences Commons [en ligne]. Disponible sur : <http://sciencecommons.org/>



: Attribution/Pas d'utilisation commerciale/Pas de modification



: Attribution/Pas d'utilisation commerciale/Partage à l'identique des conditions initiales



: Attribution/Pas d'utilisation commerciale



: Attribution/Pas de modification



: Attribution/ Partage à l'identique des conditions initiales



: Attribution

L'auteur peut également choisir d'abandonner tous les droits sur son œuvre, ce qui revient à la laisser volontairement tomber dans le domaine public, en choisissant ce symbole :



Une fois que le type de licence est choisi, l'auteur fait figurer à proximité de son œuvre cette icône qui renseigne les utilisateurs sur son choix de se placer sous le régime des Creative Commons.



L'utilisateur sait ainsi qu'il dispose de certains droits en vertu de la volonté de l'auteur³⁰. Il lui suffit de cliquer sur le symbole ci-dessus pour savoir quel type de licence a été choisi par l'auteur. S'affiche alors une fenêtre dite Common Deeds qui indique en termes clairs quelles règles s'appliquent à l'œuvre (notamment grâce aux quatre symboles de base)³¹.

L'utilisateur peut aussi cliquer sur un lien qui affiche un contrat complet exprimant en termes juridiques la teneur de la licence.

Enfin, chaque licence peut aussi être exprimé en XML, de manière à pouvoir être traitée informatiquement. Ces expressions XML des licences Creative Commons peuvent comporter de manière facultative une description de l'œuvre à laquelle elles sont liées par le biais d'éléments du Dublin Core, ainsi que des éléments portant sur la propriété intellectuelle : Createur, Contributeur et Gestion des droits, que nous avons déjà rencontrés au début de cette étude.

Les licences Creative Commons et leurs symboles sont de plus en plus utilisés par la communauté des utilisateurs Internet. Leur signification tend à être bien assimilée par les usagers. Des moteurs de recherche

³⁰ A titre d'exemple concret, les icônes qui figurent sur ce rapport sont tirés du site Internet Creative Commons. Ce site est lui-même placé sous une licence « Paternité », la moins restrictive de toute, ce qui signifie qu'il est possible de copier, de réutiliser, de transformer librement ces symboles. La seule obligation est de mentionner leur source, obligation à laquelle l'auteur de ces lignes a satisfait en rédigeant cette note.

³¹ Exemple de Common Deeds pour une licence de type Paternité/Pas d'utilisation commerciale/Pas de modification : <http://creativecommons.org/licenses/by-nc-nd/2.5>

permettent même de rechercher les œuvres qui sont placées sous licence Creative Commons (Creative Commons Search³² qui utilise Google et Yahoo).

2. Comparaison avec les besoins d'une bibliothèque numérique.

Comme les expressions XML des licences Creative Commons autorisent l'emploi du Dublin Core, on peut déjà dire qu'elles satisfont tous les besoins couverts par ce dernier langage. Mais elles permettent d'envisager d'autres possibilités, plus innovantes.

2.1. Creative Commons et œuvres tombées dans le domaine public.

On peut tout à fait concevoir qu'une bibliothèque se soucie des usages que ses clients peuvent faire des œuvres de sa collection numérique. Une question notamment est de s'assurer qu'il n'est pas fait usage des œuvres à des fins commerciales sans en demander la permission au préalable à la bibliothèque. La bibliothèque pourrait aussi souhaiter que les œuvres disponibles sur son portail ne soient pas modifiées, réutilisées, amalgamées à d'autres œuvres pour en faire des adaptations ou des œuvres composites nouvelles. La bibliothèque peut aussi vouloir exiger que l'on mentionne que l'œuvre provient de sa collection numérique chaque fois qu'elle est réutilisée, pour des raisons de visibilité. Ces demandes sont légitimes, car après tout, les collections numériques des bibliothèques publiques sont financées par de l'argent public, et même pour les œuvres qui appartiennent au domaine public, des règles de ré-utilisation peuvent être prévues.

Les Creative Commons peuvent offrir un moyen simple de manifester ces intentions à l'attention des utilisateurs à l'aide des symboles prévus par ce système.

Ainsi en plaçant les œuvres qui sont dans le domaine public sous la licence suivante :



: Attribution/Pas d'utilisation commerciale/Pas de modification

une bibliothèque obligerait à ce que son nom soit cité en cas de réutilisation et à ce que permission lui soit demandée avant toute utilisation commerciale ou modification d'une des images de sa collection.

Un seul obstacle pourrait s'opposer à cet emploi des Creative Commons pour les œuvres libres de droit de la collection. Normalement, seul le titulaire des droits peut placer son œuvre sous une telle licence, or une bibliothèque n'est pas titulaire des droits sur les œuvres qu'elle diffuse, à plus forte raison pour les œuvres tombées dans le domaine public qui sont libres de droit.

Sauf que la bibliothèque peut certainement revendiquer une forme de propriété intellectuelle, non sur les œuvres originales, mais sur les images numérisées qu'elle a produites. On peut soit considérer que la numérisation est une opération qui crée une adaptation originale de l'œuvre copiée (contestable³³), soit invoquer la protection que la jurisprudence canadienne³⁴ accorde aux créateurs de base de données. Elle est

³² Voir : <http://search.creativecommons.org/>

³³ On pourrait soutenir que la numérisation des œuvres comprend un travail de sélection, de traitement des couleurs, de contrôle de la qualité, d'ajout de métadonnées liées aux fichiers, qui autorise à y voir une œuvre de création. Mais il ne peut pas y avoir adaptation originale, à notre avis, au sens de la loi sur le droit d'auteur, car le but de l'opération est de rester le plus fidèle possible à l'œuvre qui sert de modèle. Difficile donc de voir là une forme d'originalité.

³⁴ *Tele-direct (Publications) Inc. c. American Business Information, Inc. (C.F. 1^{re} Inst.)*, conf. (1998)2 C.F. 22(C.A.F.) : une compilation de données est protégée par un droit d'auteur à condition qu'elle soit suffisamment originale, évalué à partir du degré d'habileté, de discernement et de travail consacrés à l'arrangement et à la sélection des informations. Une bibliothèque numérique structurée, organisée en corpus, accompagnée d'outils de navigation et de recherche peut tout à fait satisfaire ces critères d'originalité.

donc susceptible d'être protégée par un droit d'auteur, y compris pour la partie qui appartient au domaine public.

On voit donc l'avantage qu'une bibliothèque peut retirer à placer les œuvres tombées dans le domaine public de sa collection sous licence Creative Commons : une meilleure façon de renseigner les utilisateurs à propos des usages qui peuvent être faits de la collection, sans pour autant les entraver.

2.2 Creative Commons et œuvres protégées par des droits d'auteur

Pour les œuvres encore protégées par des droits d'auteur et pour lesquelles la bibliothèque conclue des ententes avec les ayants droit avant de les numériser et de les diffuser, il n'est a priori pas concevable de se placer sous une licence Creative Commons. Seul le titulaire initial des droits bénéficie de cette faculté.

Mais il n'est pas impossible d'envisager de demander au titulaire de se placer volontairement sous ce type de licence libre, au moment où l'on négocie l'entente avec lui. Cette solution est difficilement concevable pour les œuvres qui possèdent encore une certaine valeur commerciale. Mais pour celles qui n'en possèdent plus (et ce sont les plus nombreuses dans la collection), il n'est pas du tout impensable que les titulaires accepteraient d'adopter un des régimes des Creative Commons.

En adoptant cette solution, la bibliothèque contribuerait, par le biais de sa collection numérique à étendre le cercle des œuvres réutilisables plus librement, dans le strict respect des droits des auteurs. C'est une manière d'étendre les usages qui peuvent être fait des œuvres diffusées. Chaque auteur resterait libre de l'ampleur qu'il voudrait donner à la libération de son œuvre,

On peut aussi imaginer de créer une sous-section dans une collection numérique qui serait consacrée précisément à la diffusion d'œuvres placées sous licence Creative Commons. Une sorte de section « Œuvres Libres », séparée du reste de la collection pour éviter le mélange des genres qui peut faire peur à certains ayants droit. Un avis sur le portail de la Bibliothèque pourrait même inviter des créateurs à proposer d'eux-mêmes à la collection numérique de BanQ des œuvres placées sous licence Creative Commons. Le secteur de la Création Libre est en effet particulièrement actif et il produit des œuvres dont la qualité n'a parfois rien à envier à celle du domaine éditorial³⁵.

C'est tout de même un choix à forte valeur symbolique pour une bibliothèque publique, qui relève de la politique documentaire et pas seulement de la technique.

2.3 Creative Commons et information des usagers quant à leurs droits.

Même si les Creative Commons ne sont pas utilisés en tant que tels par une collection numérique, on pourrait s'inspirer de leur mode fonctionnement pour délivrer une meilleure information juridique aux usagers quant à leurs droits et obligations.

En général, l'information juridique délivrée aux usagers d'une collection numérique est peu visible. Il faut vraiment que l'utilisateur fasse un effort volontaire pour la trouver et la consulter. Dans ces conditions, rien ne garantit qu'un utilisateur de la collection numérique aura une idée claire de ses droits avant d'aller consulter des œuvres.

On pourrait imaginer un système similaire à celui des Creative Commons, qui affiche, en même temps que chaque œuvre consultée, un logo facilement identifiable sur lequel il suffirait de cliquer pour avoir accès à un résumé des droits et obligations de l'utilisateur. L'idéal serait que ce message soit personnalisé en fonction du statut juridique de chaque œuvre (encore protégée, dans le domaine public...). On peut également

³⁵ A titre d'exemple d'une telle bibliothèque numérique dont le contenu est libre de droit, on peut citer Wikilivres, qui regroupe des textes pédagogiques au contenu libre et gratuit. Cf. <http://fr.wikibooks.org>

prévoir un lien qui renvoie vers une page expliquant plus longuement les règles de la propriété intellectuelle pour les usagers qui désirent en savoir plus sur ces questions complexes.

En adoptant une telle « signalétique des droits », claire et frappante, nul doute que les bibliothèques publiques contribueraient à un effort de pédagogie que l'on est en droit d'attendre de leur part, surtout pour l'environnement Internet qui donne trop souvent aux usagers l'impression que « tout est permis ». Une telle signalétique peut aussi être de nature à rassurer les ayants droit sur les conditions de diffusion de leurs œuvres et faciliter les négociations.

QUATRIEME PARTIE : ODRL ET MPEG-21/5

1. DESCRIPTION

Nous avons choisi de regrouper la présentation des deux langages ODRL³⁶ et MPEG-21/5³⁷ pour plusieurs raisons.

Ce sont d'abord deux langages qui ont été prévus pour créer des environnements sécurisés permettant un contrôle effectif de la diffusion des œuvres. Ils permettent un traitement automatisé des informations qu'ils véhiculent dans le cadre d'un système de DRM. Ils sont prévus pour fonctionner en combinaison avec certains logiciels et matériels (software/hardware) qui seront capables d'interpréter les informations juridiques qu'ils diffusent pour permettre, mais surtout empêcher, certaines actions illicites comme le piratage. Ils sont donc surtout destinés à la protection d'œuvres ayant encore une valeur commerciale.

On comprend donc que ces deux langages aient été conçus par la sphère commerciale, pour des besoins qui sont avant tout ceux du commerce électronique. Cela dit, ce sont aussi les deux outils les plus puissants et les plus développés d'expression des droits disponibles actuellement. Les concepts qu'ils emploient sont par ailleurs définis de manière suffisamment générale et abstraite pour être employés en dehors de la sphère commerciale. Lorsque l'on parle de métadonnées juridiques et de DREL, il semblerait que la plupart des bibliothèques qui se posent sérieusement la question d'adopter un système examinent avant tout ODRL et MPEG-21/5. En dehors du monde des bibliothèques, ODRL et MPEG-21 sont en train de s'imposer lentement comme les deux principaux standards en matière de REL³⁸.

La grande différence entre les deux est qu'ODRL permet un contrôle poussé des usages sans l'imposer, alors que MPEG-21 est absolument dédié au contrôle de l'utilisation des ressources. ODRL est aussi un standard ouvert, à la différence de MPEG-21. Ces différences font qu'ODRL suscite plus d'intérêt auprès des bibliothécaires que MPEG-21. Il faut noter également que ces deux standards intéressent beaucoup les producteurs de E-books.

On va voir en effet que ces deux langages permettent une gestion des licences et surtout un contrôle des usages, infiniment plus riches et précis que les autres DREL précédemment évoqués. Cependant, leur origine commerciale a beaucoup orienté leur conception et il n'est pas certain qu'ils couvrent l'ensemble des besoins d'une bibliothèque numérique.

³⁶ Cf. The ODRL Initiative [**en ligne**]. Disponible sur : <http://odrl.net>. The ODRL Initiative. *ODRL : A rights expression language for digital asset management and E-Commerce* [**en ligne**], Disponible sur : <http://odrl.net/docs/ODRL-brochure.pdf>. Cover Pages. *Technology Reports : ODRL* [**en ligne**]. Disponible sur : <http://xml.coverpages.org/odrl.html>

³⁷ Organisation mondiale de normalisation. *MPEG-21 Overview v. 5* [**en ligne**]. Disponible sur : <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>. Cover Pages. *Technology Reports : MPEG Rights Expression Languages* [**en ligne**]. Disponible sur : <http://xml.coverpages.org/mpegRights.html>.

³⁸ MPEG-21/5 a même obtenu une certification ISO depuis 2004.

Ces deux systèmes de métadonnées juridiques fonctionnent de manière similaire. Le cœur du système est constitué par des informations sur les licences qui unissent des parties et définissent quels droits sont échangés relativement à une ressource donnée. ODRL et MPEG-21 permettent ensuite d'exprimer de manière extrêmement fine le contenu des licences de manière à définir exactement quels types d'usages sont autorisés, dans quelles limites et à quelles conditions.

Le plus simple pour donner une idée de la richesse de ces deux langages est de donner des exemples des termes qu'ils emploient pour identifier ces usages, ces contraintes et ces conditions.

1) Exemples d'usages

Usages	MPEG-21	ODRL
Management	delete install move uninstall	delete duplicate install move uninstall backup verify restaure
Réutilisation	adapt diminish embed enhance enlarge modify reduce	modify excerpt annotate aggregate
Transfert	transfert controll	sell lend give lease transfer
Utilisation	execute play print	display execute play print

2) Exemples de contraintes

Contraintes	MPEG-21	ODRL
utilisateur	rien de prévu	individu groupe
limites	quantity range territory	count range spatial
Temporelle	validity interval duration	datetime accumulated interval
Aspect	unit	quality unit format watermark

3) Exemples de conditions

Conditions	MPEG-21	ODRL
Païement	feeFlat feePerUse feeMetered feePerInterval	payment prepay postpay peruse
usage	seekApproval	Attribution tracked

2. Comparaison avec les besoins d'une bibliothèque numérique

2.1. Besoins en matière de description du statut juridique de la ressource et de la titularité des droits.

Développés afin de recevoir des applications commerciales avant tout, ODRL et MPEG-21 sont spécialement destinés à gérer des ressources pour lesquelles une licence a été conclue. Ils ne prennent pas vraiment en compte dans leur description les œuvres tombées dans le domaine public, ce qui peut être gênant pour une bibliothèque publique.

Les deux langages ne s'intéressent pas vraiment non plus à l'identification du titulaire des droits. Ils se focalisent sur les parties au contrat. Or celui qui accorde une licence pour une œuvre n'est pas forcément le seul titulaire des droits et il peut aussi s'agir d'une entité qui représente le titulaire, comme dans le cas d'une société de gestion.

Au final, les deux langages sont assez faibles pour ce qui est des besoins « descriptifs » d'une bibliothèque numérique.

2.2 Besoins en matière de gestion des licences et des redevances

ODRL et MPEG-21 sont spécialement conçus pour s'appliquer à des œuvres à propos desquelles des licences ont été conclues. Cependant, ils sont plus appropriés pour les cas où la licence est conclue directement entre un titulaire de droit et un utilisateur (comme par exemple entre un éditeur et un client), que pour les cas où un tiers conclue une licence avec un ayant droit pour diffuser une ressource à un public non identifiable (comme c'est le cas d'une bibliothèque numérique accessible par le biais d'Internet).

Rien n'est prévu dans ces deux langages pour recueillir des informations permettant de retracer ou de contacter les ayants droit, comme c'était le cas avec METSR.

En revanche, les deux langages incluent dans les contraintes et les conditions des éléments qui pourraient servir de support pour construire des systèmes de gestion des redevances (système de veille, de contrôle des versements..etc). Ils permettent aussi de mettre en place une tarification payable directement par l'utilisateur pour certains usages (c'est même leur but principal).

2.3 Besoins en matière de contrôle des accès et des usages.

MPEG-21 ne prévoit aucun contrôle des accès et ODRL, seulement un contrôle très rudimentaire.

C'est en effet que ces deux logiciels ont été conçus pour un environnement commercial, dans lequel on ne « filtre » pas l'accès des consommateurs aux ressources. En revanche, la sphère commerciale est très soucieuse de contrôler étroitement les usages qui sont faits des ressources qu'elle vend, et les deux langages sont particulièrement raffinés sur ce point.

La combinaison d'un très grand nombre d'éléments dans les catégories usages, contraintes et conditions permet de construire des systèmes de contraintes d'usage très élaborés. On peut en particulier imaginer un système qui prendrait en compte des catégories d'utilisateurs bénéficiant de droits d'utilisation différents en fonction des ressources. Ce genre de solution peut être utile dans une bibliothèque universitaire qui comprend des utilisateurs au statut différent.

Il est à noter également que tous les usages doivent être expressément permis avec de tels systèmes, sous peine de ne pas pouvoir exister. Cela pose des difficultés avec les œuvres libres de droit, pour lesquelles il faudrait indiquer tous les usages possibles, sous peine d'interdire des usages permis par la loi. Des systèmes automatisés sont en outre complètement incapables de gérer des notions telles que le *fair dealing*/utilisation équitable, qui par définition sont affaire d'appréciation et ne peuvent être quantifiées. Employer de tels systèmes en bibliothèque publique, c'est courir le risque de supprimer la possibilité pour les usagers de procéder à certaines utilisations couvertes par le *fair dealing*³⁹.

Au final, on se rend compte que malgré leur niveau élevé de développement, ODRL et MPEG-21 correspondent assez mal aux besoins d'une bibliothèque numérique. Les nombreux éléments qu'ils proposent sont tout de même intéressants, et pourraient inspirer les bibliothèques pour concevoir leur propre système de métadonnées juridiques.

De tels systèmes deviendraient par contre intéressants pour une bibliothèque si elle souhaitait mettre en place un environnement contrôlé autour de sa collection numérique, afin d'interdire efficacement tout usage illicite des œuvres diffusées. On pourrait même aller plus loin et contrôler de manière étroite grâce à un système de DRM mis au point grâce à ces langages l'utilisations de toutes les ressources dans la bibliothèques (CD, DVD, imprimantes, ordinateurs...etc). Mais on est en droit de se demander si des bibliothèques publiques ont intérêt à mettre en place de tels systèmes coercitifs, qui reviennent à partir du principe que les lecteurs sont tous des pirates potentiels.

La construction d'un environnement de confiance pourrait cependant être de nature à rassurer les ayants droit, qui accorderaient ainsi plus facilement leur accord pour diffuser des œuvres protégées. Ici encore, le choix de recourir à un système de DRM en bibliothèque relève plus de la politique de l'établissement que de la simple technique.

CONCLUSION : Quel système de métadonnées juridiques pour la collection numérique de BAnQ ?

Il apparaît assez nettement au terme de cette étude que les systèmes de métadonnées juridiques actuellement disponibles sont encore assez peu développés et en tout cas, généralement assez mal adaptés à l'ensemble des besoins d'une bibliothèque numérique. Aucun système ne permet par exemple de mettre en place simplement un système de veille qui indiquerait aux gestionnaires de la collection quels documents vont tomber dans le domaine public, quels documents vont nécessiter un renouvellement de licence, quels documents réclament le paiement d'une redevance pour continuer à être diffusés... En revanche, certaines métadonnées juridiques employées dans ces systèmes peuvent servir à alimenter une base de données relationnelle, à partir de laquelle on pourrait construire en interne de telles applications.

Tous les systèmes que nous avons passés en revue comportent finalement des éléments intéressants et des lacunes. Le choix d'implanter l'un plutôt que l'autre dépend étroitement des besoins de chaque bibliothèque numérique et des rapports qu'elle entretient avec son environnement.

³⁹ Dans ce sens, Coyle Karen. *Gestion des droits et besoins des bibliothèques numériques* [en ligne]. Disponible sur : http://www.nicolasmorin.com/BiblioAcid_revue/BAv1n4.pdf.

A cet égard, il serait important que les bibliothèques commencent à définir et à exprimer clairement quels sont leurs besoins en matière de métadonnées juridiques. Car en restant trop passives, elles risquent de se voir imposer des standards issus de la sphère commerciale qui ne seront pas forcément compatibles avec leurs exigences.

On peut aussi penser que les bibliothécaires pourraient être bien armés pour développer eux même des systèmes de métadonnées juridiques qui épousent aux mieux leurs besoins. Le monde des bibliothèques s'est engagé depuis plusieurs années dans la conception d'outils conceptuels très puissants, les ontologies⁴⁰, qui pourraient servir de support à de nouveaux langages d'expression des droits plus complets que ceux qui existent actuellement.

Par Lionel Maurel

Elève-Conservateur à L'Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (Villeurbanne, France)

Le 07/11/2006

⁴⁰ Les ontologies sont des modèles conceptuels qui représentent les rapports entre les grandes notions d'un secteur donné de la connaissance. Ces modèles visent à exprimer ces relations grâce à des formats informatiques qui permettront ensuite de réaliser des applications pratiques. La mise au point des ontologies informatiques constitue l'un des grands enjeux du Web 2.0 ou Web sémantique. On peut citer à titre d'exemple le modèle FRBR (Functional Requirements for Bibliographic Records) qui adapte les fonctions du catalogage classique à l'environnement numérique. Il permet en outre de décrire très finement les rapports entre une œuvre, ses adaptations, des différentes éditions jusqu'à chacun des supports sur lesquels elle est fixée. Ce modèle présente un intérêt certain pour décrire les relations de droit d'auteur sur les œuvres : Cf. IANNELLA, Renato. *Open digital rights management: a position paper for the W3C DRM Workshop* [en ligne]. Disponible sur : <<http://www.w3.org/2000/12/drm-ws/pp/iprsystems-iannella.html>>