

# Crabesques

N° 67 juillet - août - septembre 2012

## Métadonnées

de nouvelles  
convergences ?



Le fonds  
des Nations unies  
de la BIU Cujas



20 Pleins feux sur...  
Le fonds des Nations unies de la BIU Cujas



Avant le reconditionnement  
Phot. Catherine Désos-Warnier



Après  
Phot. Catherine Désos-Warnier

22 Actualité

24 Agenda

Sommaire

Dossier

**Métadonnées :  
vers de nouvelles convergences ?**

- 4 Exigences des métadonnées
- 7 Vers des référentiels globalisés
- 8 XML et les métadonnées : **irruption dans la sphère bibliographique**
- 10 Du côté d'Électre
- 11 L'enjeu des métadonnées dans l'édition et la commercialisation du livre
- 13 La réutilisation des données des éditeurs dans le cadre du dépôt légal
- 14 Un concert de métadonnées
- 17 Des identifiants normalisés à gogo
- 18 Métadonnées générées par les utilisateurs : **le cas des pratiques d'indexation sociale**

Couverture



Cerfs-volants au Cap d'Agde (8 mai 2005).  
Phot. Justmoot sur Flickr (CC BY-NC-SA 2.0)



Drapeaux des pays membres devant le siège de l'ONU à New York.  
UN Photo/Joao Araujo Pinto

# Retour sur l'année 2011

**L'année 2011 aura vu l'ouverture en rafale par l'ABES de nouveaux services et applications** : le portail des thèses (theses.fr) ; l'application STEP (Signalement des thèses en préparation), qui a pris place au sein d'un nouveau dispositif des thèses visant à mieux structurer la coordination de leur traitement par les établissements ; la signature des premiers contrats de licences nationales avec Springer, Proquest et Classiques Garnier Numérique, précurseurs du projet ISTEEX, concrétisé en 2012 et dans lequel l'ABES jouera un rôle déterminant aux côtés du CNRS, de Couperin et de l'université de Lorraine ; la version mobile du Sudoc en partenariat avec l'Université numérique Paris Île-de-France ; l'ouverture du nouveau site web de l'Agence. Un autre chantier, d'une importance stratégique pour la mise en valeur des données des bibliothèques et le positionnement de l'ABES comme fournisseur d'information bibliographique, a connu son aboutissement : l'ouverture des données du Sudoc, des thèses, de Calames et des autorités Sudoc dans le web des données. L'intégration de ces dernières dans VIAF, le fichier virtuel international des autorités géré par OCLC, participe de cette même dynamique. D'autres applications ont vu le jour, conçues pour alléger le travail des professionnels de la documentation, comme Self Sudoc, sans omettre la fin du chantier de conventionnement sur objectifs des 31 centres régionaux du Sudoc-PS qui a initié une nouvelle dynamique pour des structures parfois un peu oubliées au sein des SCD.

Parmi les chantiers ouverts en 2011, celui de la BSN est emblématique car il structure désormais l'action des opérateurs dans une volonté de rationalisation des politiques publiques. L'ABES y apporte une contribution majeure, notamment pour les segments acquisitions, signalement et modernisation du prêt entre bibliothèques aux côtés de ses partenaires, le CNRS et Couperin.

Avec le démarrage du projet de système mutualisé de gestion de bibliothèque, l'offre émergente de systèmes de nouvelle génération reposant sur l'informatique en nuage laisse présager un bouleversement du traitement de la chaîne documentaire dans les établissements et une remise en cause radicale de l'organisation du Sudoc et de son modèle actuel datant de la fin des années 90. Ce chantier est étroitement lié à celui de RDA et de FRBR sur lequel l'ABES travaille en étroite collaboration avec la Bibliothèque nationale de France.

Nous sommes entrés dans une décennie de bouleversement radical de nos outils et méthodes de travail. L'ABES est consciente que le virage doit être anticipé dès maintenant et que tous les impacts doivent en être mesurés. Son projet d'établissement 2012-2015 témoigne de ce positionnement stratégique et de cette volonté d'innovation qui s'est manifestée très concrètement par la participation réussie de l'ABES à des projets de recherche sur les processus automatisés d'amélioration de la qualité des données et catalogues : SudocAd, terminé en 2011, suivi par Qualinca à partir de 2012. Si l'ABES s'implique depuis 2011 dans le projet européen Dariah, c'est parce que l'avenir de l'information scientifique et technique ne se joue plus seulement à l'intérieur de nos frontières mais se construit avec les autres acteurs européens.

La diversification des applications et projets de l'ABES pose la question de leur mise en cohérence. Celle-ci repose sur les métadonnées, ciment de la mission et des métiers de l'Agence, qui dépassent le cadre de chacun des contextes pour se rejoindre de manière transversale. Elles requièrent des compétences multiples et des expertises simultanées sur différents plans (modèles de données, formats de saisie, normes et règles de description ; manipulation et exploitation dans les interfaces ; transformation et conversion pour l'échange avec d'autres catalogues ou applications et pour l'exposition sur le web). Le pari lancé à l'ABES est de faire converger l'ensemble de ces compétences malgré des environnements techniques hétérogènes. Ceci n'est pas sans poser des défis tant dans l'acquisition de ces compétences autour des métadonnées que dans l'organisation nécessaire à leur mise en œuvre. Faute de quoi, l'éparpillement conduirait à une balkanisation de leur approche, préjudiciable à un suivi cohérent de l'évolution des différents services et applications de l'ABES.

*Raymond Bérard*  
Directeur de l'ABES

# Exigences des métadonnées

**Q**u'une donnée puisse en représenter un jeu entier, qu'à son tour ce dernier puisse constituer un corpus cohérent de connaissances, ordonné et exploitable – cela identifie assez, depuis Alexandrie, la nécessité classificatoire des objets de pensée, leur invention<sup>1</sup> et leur mémoire.

Les métadonnées, puisqu'il s'agit bien de ces données *indexantes* du discours, ont connu une fortune invariable de renouvellement, que leur pouvoir d'orientation puis d'interprétation n'a cessé d'affirmer. Elles imposent encore aujourd'hui un maillage des contenus de la science – dans la mesure extensive de leur exposition numérique.

C'est justement sous le couvert empirique d'un univers insensé – le web de données – que la production des métadonnées trouve sa justification tant heuristique qu'économique.

## Métadonnées de propriété

Tout objet de discours (de science), toute collection de données, en définitive toute unité de sens peut recevoir une annotation minimale, une cote, un identifiant qui renvoie à un système partagé – un référentiel qui a valeur de norme et contre lequel ils sont rapportés, puis reconnus. C'est à ce premier niveau de propriété qu'une métadonnée distribue les objets comme autant d'objets uniques.

Historiquement, de tels identifiants normés ont été attribués et maintenus par des communautés d'intérêts (intellectuelles ou commerciales) exerçant un droit de propriété sur les contenus visés. Une autorité auteur par exemple, au sens bibliothéconomique du terme, peut étendre sa représentation dans une définition stricte, portée par un identifiant numérique unique. L'initiative *Orcid*<sup>2</sup> constitue à ce titre l'effort sans doute le plus significatif à l'établissement d'un registre mondial de contributeurs scientifiques, y compris d'œuvres orphelines ou posthumes. L'enjeu que représente l'établissement d'une telle métadonnée intéresse directement la propriété intellectuelle et les ayants droit d'une œuvre de l'esprit – tout en permettant de lier les objets qu'ils ordonnent à l'échelle d'une e-science dont les contributions connaissent un développement exponentiel.

Autant l'ISSN constitue une métadonnée de propriété captive – et de même le DOI qui demeure une attribution éditoriale –, autant la nécessité de métadonnées de propriété d'un autre ordre, visant tous les objets de la recherche scientifique, apparaît comme un enjeu aux nombreuses ramifications. Une extension aux données brutes de la science de ces identifiants spécifiques est notamment portée par le consortium *DataCite*<sup>3</sup>, dès lors accessibles à la citation comme aux liaisons sémantiques.

## Métadonnées descriptives

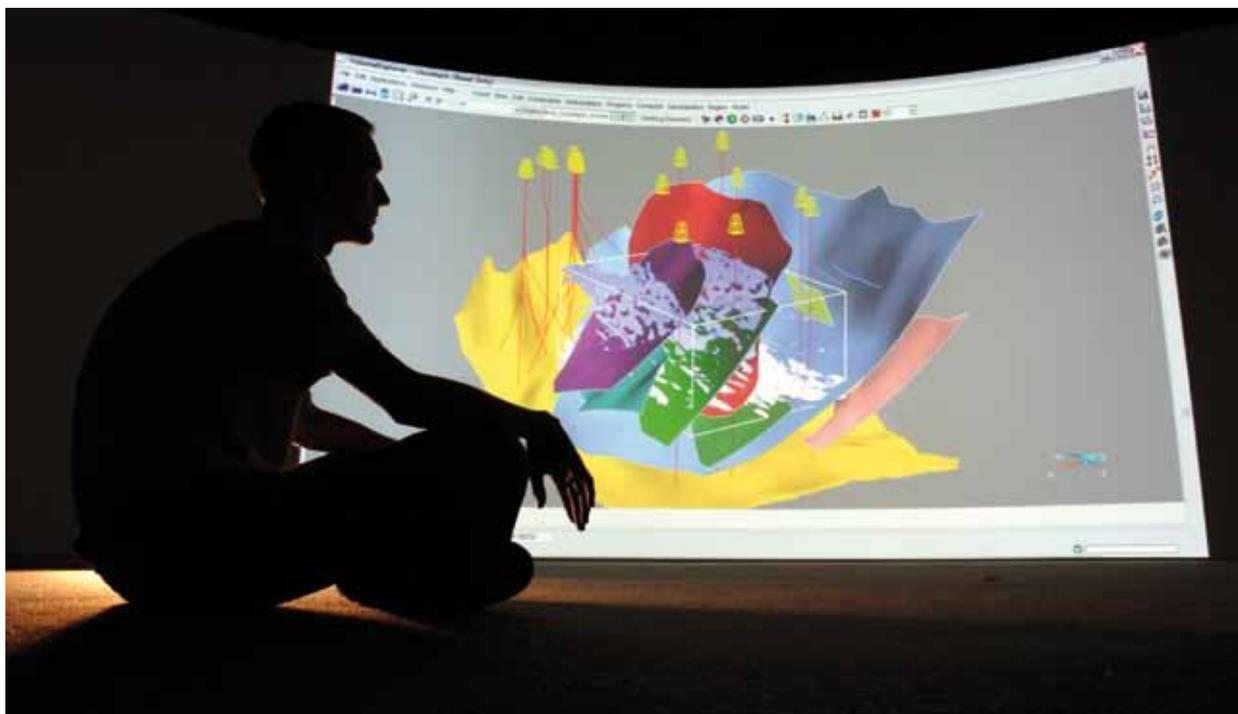
C'est une société d'informatique dans les domaines de la santé et de la médecine qui dépose la marque « Metadata » en 1986. La NASA l'utilise pour la première fois en 1988 dans une édition du *NASA's Directory Interchange Format Manual*<sup>4</sup>. Dans les années 1990, le vocable entre peu à peu dans le domaine public et s'impose comme définissant la description des données. En 1995, après qu'une première spécification ait été réalisée pour les données géographiques, le terme apparaît dans le monde bibliothéconomique avec la création du *Dublin Core Metadata Set Element*. Les professionnels de l'information prennent alors conscience qu'ils sont créateurs, depuis des décennies, de nombreuses métadonnées et que celles-ci peuvent tout autant décrire des documents papier que des ressources électroniques.

C'est ainsi que les données catalographiques, depuis leur socle descriptif MARC (*Machine Readable Cataloguing*), n'ont cessé de décliner de nouveaux modèles de représentation. Aujourd'hui, le modèle conceptuel FRBR (*Functional Requirements for Bibliographic Record*) entend formaliser les relations sémantiques entre une œuvre et ses différents types d'expressions, en s'affranchissant de la notion de « document » comme première unité signifiante. Les enjeux de ce modèle verront se démultiplier de nouveaux points d'accès à l'information et permettront de nouvelles liaisons sémantiques entre des sources de données hétérogènes.

Ces nouveaux principes, organisés autour d'entités et de relations, offriront un cadre de données structurées, appuyé sur de nombreux schémas de métadonnées XML – formats créés en fonction de l'usage des ressources<sup>5</sup>. Nous voici à l'aube de la construction du web de données, où il ne s'agit plus seulement d'assurer des liens entre des objets ou des documents, mais d'offrir un accès logique et pérenne entre les données elles-mêmes. L'accès à la connaissance suppose d'ores et déjà l'appropriation de ces nouvelles cartes du savoir, affranchi de ses limites propres.

## Métadonnées d'enrichissement

Pour porter le paradigme documentaire au sens, le seul signalement des ressources, fût-il normé et extensif, ne saurait servir qu'à un seul effet d'accumulation de données. C'est la structuration des métadonnées dans de véritables systèmes d'information, puis leurs liaisons avec d'autres univers de données, leur interopérabilité, qui permettront leur ouverture tant aux moteurs de recherche qu'aux routines des bases de connaissances. Toutes opérations qui requièrent un apport d'intelligence : le coût du facteur humain dans l'analyse et l'enrichissement des données oblige à des contournements ou à la mise en œuvre de relais sémantiques qui aient quelque valeur probante dans les différents champs de l'économie numérique.



Une première approche tient dans l'imposition de règles de production des médias d'information eux-mêmes, à des fins d'extraction automatique et contrôlée de métadonnées utiles : une telle disposition éditoriale augmente la performance des systèmes de fouille de données et des logiciels d'appariement lexicaux associés. Courante dans la chaîne éditoriale du document, de l'image ou du son (œuvre ou fragment), elle reste à généraliser aux objets complexes de la culture et de la connaissance (objets de la recherche scientifique, flux de données primaires, bruits des différents fonds de données archivées). Des ontologies y répondent, mais la méthode semble introuvable<sup>6</sup>, tant l'expression du sens d'un objet (textuel, multimédia, sériel...) semble hors de la portée des outils qui le manipulent (quand ceux-ci seraient fondés sur des recommandations du web sémantique). En clair, un jeu de métadonnées ontologiques, qui expose un contenu à l'interprétation d'agents logiciels, ne saurait épuiser le sens qu'il porte, sauf à le rapporter à un segment de science (une discipline captée par son lexique) ou de culture (une œuvre représentée par ses codes) – c'est-à-dire à des référentiels fermés, sans prétention herméneutique (interprétative)<sup>7</sup>.

Mais à défaut de sens, ces métadonnées référentielles peuvent opérer une aire de correspondances entre différents univers de données. Des objets liés à des référentiels hétérogènes peuvent entrer en résonance, sans rompre leurs lexiques propres. Des initiatives comme SKOS cherchent les conditions d'interopérabilité de ces différents référentiels auxquelles de grands prescripteurs institutionnels apportent leur soutien<sup>8</sup>.

## Une exposition au plus grand nombre

Un des enjeux d'exploitation de la variété des gisements informationnels repose sur leur facilité d'exposition (*Open Data*<sup>9</sup>) et d'intégration dans des systèmes tiers, au-delà de leurs éventuelles incompatibilités natives pouvant résulter de leur structuration physique ou sémantique.

À titre d'exemple, la complexité de la description de l'objet muséal ou archéologique, dans toute la variété de ses

formes, a conduit à la création de modèles, tel CRM-CIDOC<sup>10</sup> devenu norme ISO en 2006, et à plusieurs standards pour les décrire tels CDWA (*Categories for Description of Work of Art*)<sup>11</sup> et LIDO<sup>12</sup>.

CRM-CIDOC, ontologie de domaine orientée objet, entend modéliser et formaliser la nature et les relations existant entre divers éléments descriptifs des objets de musée, qu'ils soient unitaires ou partie d'un tout, réels ou virtuels. Ce modèle s'accorde à donner le sens précis des métadonnées ainsi que des événements entrant en relation avec l'objet décrit. Ces descriptions fortement contextualisées permettent de préciser l'unicité de l'item (en restituant ses identifiants, ses marques, ses mesures...) tout en offrant des liens d'adjacence avec des ressources externes qui qualifient son environnement (géographique, historique, manifestations en rapport, intervenants...).

Appliqué à divers corpus documentaires, tels qu'images, textes, vidéos, CRM-CIDOC peut offrir de nouveaux espaces de médiation fédérant des ressources variées et favorisant ainsi la découverte, au-delà des frontières disciplinaires ou institutionnelles. L'initiative CLAROS (*CLassical Art Research Online Service*)<sup>13</sup>, programme soutenu par l'UNESCO, réunit plus de 20 millions de notices d'objets de l'Antiquité gréco-romaine issus des musées du monde entier, grâce à l'utilisation du CRM-CIDOC et de RDF, le tout interrogeable *via* le langage de requête SPARQL.

La création de ces nouveaux points d'accès permet une plus vaste exposition des métadonnées, véritable enjeu de diversité culturelle et d'aide à la recherche scientifique, rendu possible grâce aux langages et vocabulaires sémantiques tels que RDF, SKOS...

Dans un contexte plus commercial, les métadonnées d'identification, telles ISAN (*International Standard Audiovisual Number*), ISRC (*International Standard Recording Code*) et ISMN (*International Standard Music Number*), prennent tout leur sens dans un contexte de reproductibilité des supports, de quantité de transactions et d'échanges dans le cadre mondial du e-commerce. En tant qu'identifiants uniques, ces formats répondent à une attente très forte de la part des ayants droit, des gestionnaires de col-

lections, des diffuseurs ou revendeurs de contenus. En général, ils sont attribués par une autorité fédérative gérant un répertoire central.

## Et demain ?

Ces métadonnées, embarquées dans de nouveaux médias, utilisées *via* des dispositifs ubiquitaires (extensifs et mobiles), se proposeront demain spontanément aux utilisateurs que nous sommes, en fonction de nos usages passés – appréciations, prescriptions et consommations. Alors qu'à ce jour, l'utilisateur numérique se devait d'être actif pour accéder à l'information, celle-ci pourrait être dotée d'une forme d'intelligence afin d'anticiper les besoins informationnels, préparant ainsi le web de demain : l'Internet des objets<sup>14</sup>.

Jacques Millet  
Gaëlle Rivérieux

Inria, établissement public de recherche dédié aux sciences du numérique <http://www.inria.fr>  
Jacques Millet est délégué à l'information scientifique de l'Inria et membre du Comité scientifique de l'ABES [jacques.millet@inria.fr](mailto:jacques.millet@inria.fr)  
Gaëlle Rivérieux est responsable IST du Centre Inria Grenoble Rhône-Alpes [gaelle.riverieux@inria.fr](mailto:gaelle.riverieux@inria.fr)

<sup>1</sup> *Inventio* – au sens latin de découverte.

<sup>2</sup> *Open Researcher and Contributor ID* – proposé comme identifiant unique et ouvert pour les contributeurs d'œuvres de l'esprit, similaire à ce que pourrait être un DOI [*Digital Object Identifier*] d'auteur (cf. <http://www.orcid.org>). Voir également, sur un périmètre élargi des œuvres de création, l'ISNI [*International Standard Name Identifier*] qui ambitionne l'interopérabilité de systèmes d'identification propriétaires.

<sup>3</sup> Datacite.org – et son *Metadata Store* : <https://mds.datacite.org/>

<sup>4</sup> Caplan (P.), *Metadata for all librarians*, Chicago, American Library Association, 2003, p. 1.

<sup>5</sup> Dublin Core, EAD, MODS, METS...

<sup>6</sup> C'est cependant l'horizon que se donne RDF [*Resource Description Framework*], avec l'établissement de représentations sémantiques des métadonnées descriptives des objets de recherche et de culture – sans le filtre d'une lecture humaine.

<sup>7</sup> Cf. Zacklad (M.), « Évaluation des systèmes d'organisation des connaissances », *Les Cahiers du numérique*, 2010, vol. 6, n° 3, p. 133-166.

<sup>8</sup> Bibliothèques nationales, ministères de la Culture, registres de publications officielles.

Voir <http://www.w3.org/2001/sw/wiki/SKOS/Datasets>

<sup>9</sup> <http://www.data.gouv.fr>

<sup>10</sup> <http://www.cidoc-crm.org>

<sup>11</sup> Développé au début des années 1990 : [http://www.getty.edu/research/publications/electronic\\_publications/cdwa/index.html](http://www.getty.edu/research/publications/electronic_publications/cdwa/index.html)

<sup>12</sup> <http://www.lido-schema.org/documents/LIDO-Handout.pdf> (2010).

<sup>13</sup> <http://www.clarosnet.org>

<sup>14</sup> Mitton (N.), Simplot-Ryl (D.),

*From the Internet of things to the Internet of physical world*, <http://hal.archives-ouvertes.fr/hal-00598395>

## Les métadonnées, ces grandes indiscretes

Chacun a encore en mémoire l'affiche de campagne de Nicolas Sarkozy pour les dernières élections présidentielles : le portrait du candidat avec le slogan « Une France forte » au premier plan, une mer calme et un ciel lumineux, au second.

Patrick Peccatte, chercheur associé au Laboratoire d'histoire visuelle contemporaine (Lhivic/EHESS), dans un billet publié le 17 février 2012 sur Culture visuelle, explique comment un internaute a réussi à établir que la photo du paysage marin de l'affiche n'était autre qu'un cliché de la mer Égée en Grèce.

C'est la diffusion de l'affiche au format numérique qui a permis de remonter ainsi jusqu'à la prise de vue initiale.

L'auteur du billet rappelle les 3 types de métadonnées stockées dans un fichier image : **les informations IPTC/IMM** (*Information Interchange Model*), standard adopté en 1994 par Adobe pour définir les informations associées à une image et dont les champs sont renseignés manuellement lors de son indexation ; **les données EXIF** (*EXchangeable Image File*), informations d'ordre technique correspondant à un standard supporté par tous les fabricants d'appareils numériques, renseignées automatiquement lors de la prise de vue ; **les données XMP** (*eXtensible Metadata Platform*), créées par Adobe en 2001, version simplifiée de RDF (*Resource Description Framework*), modélisation de base du web sémantique.

Le croisement de ces trois types de métadonnées a permis de retrouver l'agence qui a commercialisé la photo (Tetra Images) et d'établir que cette dernière a été prise en Grèce.

Des investigations un peu plus approfondies fournissent la date de prise de vue du cliché initial (28 mai 2011) et l'examen précis des données XMP de l'affiche numérique dévoile que la photo a subi 38 modifications du 12 au 14 février 2012.

La révélation de ces investigations sur les métadonnées de l'affiche a bien évidemment donné lieu à un *buzz* sur Internet pointant la maladresse du slogan « une France forte » associé à une photo de la Grèce, pays dont on connaît la situation économique actuelle. Un *buzz* dont les concepteurs de l'affiche se seraient sans doute fort bien passés mais qui prouve que désormais les communicants devront apprendre à se méfier des indiscretions des métadonnées.

Béatrice Pedot

Pour en savoir plus : <http://culturevisuelle.org/dejavu/1118>

# Vers des référentiels globalisés

**C**omment identifier de manière précise et durable Raymonde Moulin, le concept de « moulin » ou la ville de Moulins ? Dans notre monde à nous, celui de l'information bibliographique, on s'appuie sur nos propres référentiels, historiquement liés au Sudoc : fichier d'autorité de noms de personne, vocabulaire contrôlé Rameau ou MeSH, fichier d'autorité de noms géographiques, etc. Dans le périmètre de l'enseignement supérieur et de la recherche en France, on a souhaité donner à ces référentiels une autonomie fonctionnelle et technique, afin de permettre à d'autres applications que le Sudoc d'utiliser et d'alimenter IdRef, comme il faut désormais appeler ces « autorités Sudoc ».

## Des référentiels plus ouverts

Le mouvement d'ouverture d'IdRef vers d'autres applications illustre à l'échelle nationale et universitaire un mouvement global d'ouverture des référentiels. Dans le monde des bibliothèques, il faut citer VIAF, fichier d'autorité international virtuel – virtuel au sens où il est l'agrégation et surtout l'interconnexion de plusieurs fichiers d'autorité

nationaux, dont celui de la Bibliothèque nationale de France et IdRef. Mais VIAF est lui-même agrégé et interconnecté dans une base plus large et moins spécifiquement bibliographique : ISNI (norme ISO ; *International Standard Name Identifier*). Cette tendance accompagne et enrichit la croissance du web et notamment du web de données, avide d'identifiants globaux pour référencer et décrire de manière univoque toutes sortes d'entités : personnes, organismes, concepts, lieux, époques, microbes, émotions, événements, langues, etc. Tout ce dont on peut parler, il faut l'identifier, au moyen d'un code (un numéro PPN dans IdRef ; une URL sur le web). Le nouveau contexte est global en deux sens : il dépasse les frontières locales ou nationales ; il dépasse les frontières professionnelles (le bibliothécaire a besoin de référencer sans ambiguïté la ville de Moulins, mais c'est aussi le cas de l'administra-

tion, des offices du tourisme, des épidémiologues, des biographes de Georges Bidault et même des indépendantistes bourbonnais). Très récemment, Google et d'autres moteurs de recherche majeurs ont investi la question des référentiels, en mettant sur la table schema.org<sup>1</sup>. Jusqu'à présent, le web de données se construisait de manière relativement décentralisée, grâce à la rencontre de diverses initiatives, de vocabulaires de description multiples et de référentiels variés. Avec schema.org, Google et ses alliés entendent remettre de l'ordre dans cette auberge espagnole, en imposant un vocabulaire unique et en adoubant, voire en absorbant, certains référentiels – et pas d'autres<sup>2</sup>.

## Faire cohabiter des référentiels locaux et globaux

Avec schema.org, la mondialisation des métadonnées risque de prendre une tournure unilatérale, appauvrissant de ce fait la richesse des métadonnées et donc les métadonnées elles-mêmes... à moins qu'on en fasse un bon usage. Car il faut plutôt voir dans schema.org (et même VIAF) une *lingua franca*, outil pratique de communication qui ne doit pas se substituer aux langues vernaculaires plus expressives et enracinées dans des pratiques professionnelles voire culturelles – et non décrétées depuis Mountain View, CA. Par exemple, nous devons conserver la maîtrise d'un référentiel de personnes physiques et de collectivités constitué et entretenu à des fins bibliographiques et universitaires, même s'il faut se brancher sur le niveau global en liant IdRef à VIAF ou DBpedia (Wikipedia version web de données). Au contraire, nous avons pris l'habitude d'utiliser directement les normes ISO qui référencent les pays ou les langues – en tolérant au passage certains sacrifices délicats.

Il n'est pas toujours facile de choisir : s'il n'est pas question de gérer dans un référentiel propre à notre métier, comme Rameau ou même MeSH, la diversité et le volume d'information sur la drosophile que contient Flybase, on peut se poser la question pour identifier des lieux (noms géographiques), des marques commerciales, des genres littéraires, des concepts mathématiques, etc. Il s'agit là de faire des choix stratégiques de long terme, en prenant en considération des facteurs fonctionnels (structure et finesse de l'information, interopérabilité) et économiques (capacité à maintenir collectivement un référentiel, viabilité des organisations extérieures chargées de maintenir un référentiel).

Yann Nicolas

Chef du département Études et Projets de l'ABES

[nicolas@abes.fr](mailto:nicolas@abes.fr)

<sup>1</sup> <http://schema.org>

<sup>2</sup> <http://www.w3.org/wiki/WebSchemas/ExternalEnumerations>



Le Vieux Moulin(s) à Moulins (03).  
Phot. Thomas Durand sur Wikimedia  
Commons (CC BY-SA 3.0)

# XML et les métadonnées

## Irruption dans la sphère bibliographique

« Rien ne se perd, rien ne se crée, tout se transforme » (Antoine Lavoisier)

**S**i on prend pour angle d'observation l'essor fulgurant des grandes bibliothèques numériques, comme Gallica, Europeana ou Trove – pour ne citer qu'elles –, on perçoit d'emblée que la voie d'un nouveau fondamental du signalement bibliographique est ouverte. Afin de garantir l'interopérabilité, non plus seulement entre les architectures techniques, mais aussi entre les formats de données, l'adoption de nouvelles règles et de bonnes pratiques communes s'est avérée indispensable. Clé de réussite pour une véritable intégration de l'ensemble des ressources numériques au sein des catalogues de bibliothèque, c'est sous les auspices non plus du catalogage, mais du *Metadata Management* que se profile cette (r)évolution.

### Focus sur la courte histoire des métadonnées

C'est tôt dans l'histoire du web, lors de la conférence WWW Geneva 94<sup>1</sup> au cours de laquelle fut par ailleurs annoncée la création du W3C, que la notion de métadonnées (en anglais : *Metadata*) est proposée par l'inventeur du web, Tim Berners-Lee. Données signifiantes destinées à faciliter l'accès au contenu informationnel d'une ressource numérique, les métadonnées sont en effet au cœur de l'architecture du web – et, de fait, au cœur du métier de bibliothécaire !

L'ISO les caractérise comme des « données qui définissent et décrivent d'autres données » et on peut, à ce titre, considérer les données bibliographiques au format ISBD ou leur traduction aux formats MARC comme des « métadonnées (externes) qui s'ignorent » !

Mais qu'il s'agisse de notices décrivant un document (métadonnées externes), de marqueurs introduits directement dans un fichier informatique (métadonnées encapsulées), ou de schémas prédéfinis pour intégrer tout à la fois le document lui-même et les jeux de métadonnées qui lui sont associés (métadonnées englobantes), les métadonnées ont un rôle commun : elles permettent le déploiement d'un système d'information apte à répondre à l'ensemble des besoins inhérents aux ressources numériques (identification, description, administration, gestion technique, des droits, de l'archivage pérenne...).

Autre dénominateur commun : le langage XML, grâce auquel les métadonnées ont trouvé un moyen d'expression privilégié, à tel point que l'une de leurs caractéristiques est justement d'être « encodées en XML ». En effet, ce langage extensible, standardisé par le W3C, répond aux besoins d'échanges de données entre systèmes d'informations hétérogènes. On dit communément d'XML qu'il est un langage pivot, par définition à l'aise dans la jungle des formats : indépendant car non propriétaire, il répond par-

faitement à la nécessaire interopérabilité entre les matériels, les logiciels, les structures de données, les interfaces. Simple fichier texte décrivant des données en séparant la présentation et le contenu, destiné à être interprété par une machine (en anglais : *Machine Readable*), tout en restant lisible par un être humain, XML, de par l'extensibilité de sa structure, a l'avantage de pouvoir gérer l'arborescence des données et la notion d'héritage. Ainsi, en s'insérant dans des structures prédéfinies, qu'il s'agisse des DTD (*Document Type Definition*), vocabulaires définissant les modèles, ou des schémas, structures destinées à englober l'ensemble des jeux de métadonnées et contenus associés, les métadonnées écrites en XML peuvent rester conformes aux normes et formats « métiers », garants d'une certaine cohérence en termes de normalisation nationale et internationale, tout en adaptant leur structure pour définir les différents niveaux de granularité d'information et permettre leur portabilité.

Concernant les formats propres aux notices bibliographiques, le développement d'une logique reposant sur les métadonnées a mis très tôt en évidence le manque flagrant d'un modèle conceptuel auquel se référer : ce constat est à la base des « Spécifications fonctionnelles des notices bibliographiques » (en anglais : *Functional Requirements for Bibliographic Records*, FRBR), modélisation des informations contenues dans les notices bibliographiques, développées par un groupe d'experts de l'IFLA de 1991 à 1997, approuvées en 1997 par le Comité permanent de la Section de catalogage de l'IFLA, et publiées en 1998. Complété par un modèle analogue pour les données d'autorité (FRAD<sup>2</sup>), FRBR constitue l'ossature des nouveaux codes de catalogage, au premier rang desquels RDA (Ressources : description et accès), qui se positionne d'ores et déjà comme le futur code international.

### La mutualisation pour une plus grande dissémination

Pour définir les schémas de métadonnées indispensables au développement des bibliothèques numériques et à l'adaptation des catalogues, la Bibliothèque du Congrès a joué un rôle moteur en élaborant, dès 1999, le schéma MARC-XML, pour permettre la conversion du format MARC21 (et consorts) en XML, puis les schémas MODS<sup>3</sup> et MADS<sup>4</sup>, dédié aux autorités, et enfin, le schéma METS<sup>5</sup>, format d'implémentation permettant de combiner différents jeux de métadonnées descriptives, techniques, de droits et de préservation. Avec METS, sur le modèle duquel sont encodées par exemple les données de la bibliothèque numérique Gallica, c'est l'ensemble des besoins de la vie du document numérisé qui est couvert.



Au niveau national, l'implémentation des schémas de métadonnées a également été mise en œuvre depuis quelques années, notamment pour les outils de signalement mutualisés gérés par l'ABES : on citera par exemple l'EAD (*Encoded Archival Description*) ou TEF<sup>6</sup>, utilisés respectivement pour encoder les données patrimoniales dans le cadre de Calames (Catalogue en ligne pour les archives et manuscrits de l'enseignement supérieur), et celles des thèses électroniques via STAR (Signalement des thèses électroniques, archivage et recherche), outil sur lequel repose en partie « theses.fr », brique majeure pour la valorisation des thèses françaises.

Indépendamment de l'évolution des formats proprement bibliographiques, le Dublin Core est devenu « l'autre » schéma de métadonnées incontournable pour la gestion des ressources électroniques. Commandité par OCLC et NCSA (*National Center for Supercomputing Applications*), le Dublin Core tire son nom du groupe de travail qui s'est réuni en 1995, à Dublin (Ohio). Il s'agissait pour ce groupe d'experts issus d'univers professionnels divers, de définir un tronc commun élémentaire, utilisable pour la description de toutes ressources numériques. Schéma de métadonnées dites encapsulées, le Dublin Core comporte 15 éléments répartis au sein de 3 domaines : le *contenu* (titre, sujet, description, source, langue, relation, couverture) ; la *propriété intellectuelle* (créateur, éditeur, contributeur, droits) ; la *matérialisation* (date, type, format, identifiant). Par sa simplicité, le Dublin Core est apparu d'emblée comme le schéma le plus efficace pour structurer les entrepôts de métadonnées communs aux archives ouvertes, aux musées, aux archives audiovisuelles, au patrimoine écrit numérisé, ressources (jusqu'alors) décrites dans des formats spécifiques propres à chacune des institutions et, par définition, difficilement interopérables.

Le Dublin Core s'est en outre imposé comme le moyen le plus efficace dans le cadre notamment des programmes du mouvement *Open Access Initiative*, au travers par exemple du protocole OAI-PMH (*Open Archive Initiative - Protocol for Metadata Harvesting*). Le principe est simple : il s'agit de mettre à disposition les métadonnées concernées, en les exposant au sein d'entrepôts afin de les rendre récupérables par des moissonneurs. C'est de cette façon que les grandes bibliothèques numériques s'alimentent en données extérieures, multipliant les points d'accès (dissémination) et garantissant ainsi une plus grande diffusion. À titre d'exemple, nous citerons Gallica (qui moissonne les métadonnées exposées par une quinzaine de bibliothèques patrimoniales françaises), Europeana (qui moissonne les données des bibliothèques nationales européennes), mais aussi Dart Europe<sup>7</sup> (point d'accès vers les thèses européennes disponibles en texte intégral qui moissonne, par exemple, les données de TEL<sup>8</sup> et de STAR, entre autres dépôts institutionnels dédiés aux thèses).



#### **Mutualiser pour mieux disséminer.**

Phot. Barcoseb sur Flickr (CC BY-NC-ND 2.0)

À l'ère du *Metadata Management*, les opérateurs nationaux et internationaux (agences bibliographiques, bibliothèques nationales, consortia) se situent donc, de fait, en tant que « hub de métadonnées », élément phare du nouveau projet d'établissement de l'ABES.

Christine Fleury  
Conservatrice à l'ABES  
✉ [fleury@abes.fr](mailto:fleury@abes.fr)

<sup>1</sup> Extrait de la conférence:

✉ <http://www.w3.org/Talks/WWW94Tim/>

<sup>2</sup> FRAD : *Functional Requirements for Authority Data*, 2009.

<sup>3</sup> MODS : *Metadata Object Description Schema*, publié en 2003 ; il s'agit d'une simplification de MARC 21.

<sup>4</sup> MADS : *Metadata Authority Description Schema*, publié en 2004.

<sup>5</sup> METS : *Metadata Encoding and Transmission Standard*, publié en 2001.

<sup>6</sup> TEF : Thèses électroniques françaises - Recommandation issue des travaux de l'AFNOR (CG46 /CN357/GE5) :  
✉ <http://www.abes.fr/abes/documents/tef/>

<sup>7</sup> Dart Europe : ✉ <http://www.dart-europe.eu>

<sup>8</sup> TEL : Thèses en ligne - ✉ <http://tel.archives-ouvertes.fr/>

## Du côté d'Électre

### Les métadonnées, vitrines du web, et nouveaux enjeux

Avec la dématérialisation croissante des contenus liés aux nouveaux usages du numérique, les organismes travaillant sur les normes et les standards sont de plus en plus sollicités.

De nouvelles logiques liées au concept de dématérialisation ont fait apparaître des notions abstraites comme l'œuvre qui permet de regrouper l'ensemble de ses manifestations sur tous types de médias et de supports.

Électre s'est associé à la société Antidot pour faire une démonstration de faisabilité autour du web sémantique, durant le dernier Salon du livre, en croisant plusieurs sources, comme **IMDb**, **Freebase**, **DBpedia**, **MusicBrainz**, afin de créer un nouvel univers où se mêlent adaptations cinématographiques, livres et musiques de films. La route structurée du web se globalise et les données s'y publient directement.

L'orientation de la production bibliographique d'Électre, depuis sa première informatisation en 1984, a toujours été de faire un catalogage enrichi, notamment du résumé, pour tous types d'ouvrages.

Avec la production rendue visible sur *Livres-Hebdo*, il a fallu catégoriser à la fois pour les éditeurs, les libraires, les bibliothèques et centres de documentation. Ces besoins nous ont conduits à enrichir qualitativement nos données en ciblant ces professions.

Aujourd'hui, en sus du catalogage, nous trouvons dans une notice pivot XML Électre : un résumé ; la première et la quatrième de couverture ; une table des matières ; le feuilletage de l'ouvrage ; les indexations Rameau, Dewey et Lexique (thésaurus métiers) ; un thème à trois niveaux d'arborescence ; un genre ; un public ou niveau de lecture et enfin une thématique par clients.

L'ensemble des éléments en provenance de notre système de production est mis dans un modèle orienté autour de l'œuvre et l'historique de la base a été rétroconverti dans cette modélisation. Nous ne parlons plus de notices, mais d'un système en flux de production qui peut s'orienter vers du papier, de l'écran ou du web. Ceci permet de fournir, en plus de la donnée, l'usage ou le service contextualisé.

La base Électre est dorénavant plus utilisée par des applications web que par sa consultation seule.

Vers un XML sémantique ? Le groupe **EDITEUR** mène une réflexion sur une modélisation objet unifié (UML) de la norme Onix, début d'une conceptualisation virtuelle pour de nouveaux usages, notamment sémantiques.

De futurs modèles complexes, de plus en plus interconnectés, vont engendrer de nouveaux enjeux économiques et sociaux. Ces écosystèmes devront être réalisés avec le plus grand soin dans le respect des normes et des standards développés par les structures internationales telles **NISO**, **W3C**, **EDITEUR**... C'est dans ce cadre que le Cercle de la Librairie a été désigné comme référent Onix en France.

*Laurent Dervieu*

Responsable produit Électre

[ldervieu@electre.com](mailto:ldervieu@electre.com)

#### Qu'est-ce que c'est ?

**IMDb (Internet Movie Database)** : base de données sur le cinéma mondial, la télévision et les jeux vidéo. Propriété d'Amazon.com, son enrichissement et sa mise à jour sont ouverts gratuitement à tous les internautes. <http://www.imbd.com>

**Freebase** : projet collaboratif de rassemblement des connaissances sous forme sémantique, distribué sous licence Creative Commons par la société Metaweb. <http://www.freebase.com>

**DBpedia** : projet d'extraction de données de Wikipedia pour une version web sémantique structurée, mené par l'université de Leipzig, l'université libre de Berlin et Openlink Software. <http://dbpedia.org>

**Musicbrainz** : base de données musicales et librement diffusables, hébergée par Metabrainz, qui référence des enregistrements d'œuvres. <http://musicbrainz.org>

**EDITEUR** : organisme international de coordination du développement et de la promotion de standards pour le commerce électronique du livre. <http://www.editeur.org>

**W3C** : consortium international qui développe des standards afin de faciliter la compatibilité des différentes technologies du web. <http://www.w3.org>

**NISO** : organisme de normalisation américain consacré à la documentation et aux livres (à l'origine de la fameuse norme Z 39-50). <http://www.niso.org>

# L'enjeu des métadonnées

## dans l'édition et la commercialisation du livre

**D**epuis le début des années 2000, les éditeurs découvrent un nouveau métier : celui de... bibliothécaire. Ou, plus précisément, l'une de ses plus anciennes missions, le catalogage des livres. Jusqu'alors, les éditeurs ignoraient largement les techniques de production des données techniques, commerciales et bibliographiques servant à décrire les livres qu'ils publiaient – et *a fortiori* des enjeux liés à cette production. Mais l'essor spectaculaire des « librairies en ligne » (plus de 10 % du marché en 2011, souvent beaucoup plus pour certains secteurs), puis le développement de l'édition électronique depuis 2007, ont totalement changé la donne, obligeant les éditeurs à intégrer largement chez eux la production de ces fameuses « métadonnées ». Pour prendre la mesure de ce bouleversement, un bref historique (limité au cas français) n'est pas inutile.

### Des bases de données conçues sans les éditeurs...

Depuis 1856, le Cercle de la librairie, réunissant libraires, éditeurs et imprimeurs, publiait la *Bibliographie de la France* (créée en 1811), dont la partie proprement bibliographique deviendra dans les années 1970 une publication annuelle, *Les Livres disponibles*. Au début des années 1980, la création par le Cercle de la base de données Électre (lancée en 1986), destinée à rendre ces données bibliographiques accessibles par voie informatique aux professionnels (libraires et bibliothécaires) comme aux particuliers (via le Minitel), se fit pour l'essentiel sans la participation des éditeurs, lesquels avaient déjà mis quelques années à se convertir à l'usage de l'ISBN, introduit en 1972. Électre fut d'ailleurs, dans un premier temps, un fiasco technique et financier, qui ne fut surmonté que par le soutien décisif du directeur du livre et de la lecture Jean Gattégno, puis par la scission en 1988 de la société – « Électre Bibliographie » (l'actuelle Électre) assurant les fonctions de catalogage et « Électre Transmission » (devenue ensuite Édillectre puis Dilicom, cogérée par les distributeurs et les libraires) celles de transmission par échange de données informatisé (EDI) des commandes des libraires aux distributeurs.

Au fil des années, ces deux structures interprofessionnelles sont devenues remarquablement efficaces, mais cela s'est fait pour l'essentiel sans les éditeurs. Ceux-ci se sont contentés de sous-traiter à leurs distributeurs la mise en forme et la (complexe) normalisation des données techniques et commerciales de leurs livres, via Dilicom et la Commission de liaison interprofessionnelle du livre (CLIL, constituée en 1991 pour gérer le transport du livre), qui ont créé en 2001 (avec Hachette) le « Fichier exhaustif du livre » (FEL), lequel comptait plus de 1 100 000 références en 2009. Tandis qu'Électre continuait à créer des notices

bibliographiques au format UNIMARC (1 200 000 titres en 2012, dont 500 000 épuisés et 14 000 à paraître), à partir des spécimens envoyés par les éditeurs, puis de leurs transferts par EDI de quelques métadonnées basiques.

En bref, un travail faisant largement double emploi avec celui du Catalogue général de la BNF (constitué à partir du dépôt légal), devenu lui aussi de plus en plus performant dans les années 2000, au point de concurrencer désormais Électre dans l'établissement rapide de notices bibliographiques des nouveautés. Sans parler des multiples autres bases de données bibliographiques et commerciales créées depuis les années 1990 par divers acteurs privés (comme Tite-Live ou Decitre) pour pallier les insuffisances des outils institutionnels (tarifs trop élevés d'abonnement à Électre, délais jugés trop longs de la BNF, etc.).

Depuis plus de vingt ans, nous avons été quelques-uns, bibliothécaires, libraires et éditeurs, à tenter de faire converger ces entreprises inutilement concurrentes, mais sans succès. Un exemple parmi d'autres : dans les années 1990, afin de faciliter le travail des bibliothécaires, les éditeurs de sciences humaines et sociales du groupe universitaire du Syndicat national de l'édition<sup>1</sup> (SNE) ont obtenu d'Électre qu'elle produise à l'avance des notices de catalogage (avec références Rameau et Dewey) qu'ils faisaient figurer sur la page de copyright de leurs livres, comme cela se faisait déjà de longue date dans nombre d'autres pays, des États-Unis au Brésil. Mais les responsables d'Électre de l'époque ont finalement renoncé à produire ce service, jugé trop coûteux...

Ce rappel historique indique à quel point la plupart des éditeurs français sont restés longtemps étrangers à la culture de la production de métadonnées, au mieux considérée comme une affaire de commerçants (les libraires, à faire traiter par leur distributeur) ou de conservateurs (les bibliothécaires). C'est surtout l'émergence imprévue du « commerce en ligne », à partir de 2000, qui les a amenés à se remettre en cause sur ce point.

### ...à la production de métadonnées par les éditeurs...

Les grands groupes d'édition, parce qu'ils avaient les moyens de ce lourd investissement, ont été les premiers à créer des bases de données commerciales et bibliographiques internes, directement renseignées par les éditeurs eux-mêmes, à partir de formulaires précis et normalisés. À La Découverte, nous avons ainsi été partie prenante du projet « Yod@ », lancé en 2001 par notre groupe Vivendi Universal Publishing (devenu Editis en 2004). Celui-ci consistait à créer, pour chaque livre à paraître, une fiche renseignant tous les champs pertinents pour le décrire : titre et sous-titre (provisaires puis définitifs), auteurs, argumentaire puis quatrième de couverture, image de cou-

verture, table des matières, pagination, prix, format, catégorie CLIL et des dizaines d'autres champs actualisés au fil de la vie du livre (avant et après parution).

Cela nous a conduits à revoir en profondeur nos procédures de fonctionnement interne, facilitant la vie de tous les collaborateurs, puisque les données d'un livre n'étaient plus saisies qu'une seule fois – c'était loin d'être le cas auparavant – et devenaient accessibles à tous, par exemple pour produire de façon automatisée nos catalogues imprimés ou pour alimenter en permanence notre site web, créé en 2002 (lequel, de plus en plus visité, nous a permis d'abandonner en 2007 notre bulletin trimestriel imprimé annonçant nos nouveautés, remplacé par une *newsletter* électronique).

L'amélioration progressive de cette base, étendue à l'ensemble des catalogues des éditeurs d'Editis, a permis à notre distributeur Interforum de la mettre progressivement à disposition (gratuitement) des libraires en ligne. Non sans maints « bugs » et difficultés, faute d'une normalisation interprofessionnelle des protocoles techniques de production de ces métadonnées. La norme Onix<sup>2</sup> française élaborée pour résoudre ce problème (à partir de son modèle d'origine états-unien), dont les prolégomènes avaient été pourtant engagés à l'initiative du Cercle de la librairie dès 2001, a mis en effet de longues années à s'imposer. Même si la « commission numérique » de la CLIL, créée en 2011, œuvre désormais pour généraliser cette norme, tous les grands diffuseurs/distributeurs français n'ont pas achevé en 2011 cette migration (Interforum l'a fait en 2011, sa base Yod@ devenant « Yonix »). Et les plus petits diffuseurs/distributeurs, au service de centaines d'éditeurs plus « artisanaux », restent encore derrière, malgré leurs efforts constants.

Car cette fois, on ne peut plus incriminer le désintérêt des éditeurs face à l'importance de produire, en même temps que les livres qu'ils publient, des métadonnées numériques associées de bonne qualité. La grande majorité d'entre eux, quelle que soit leur taille, ont bien compris l'importance de cet enjeu, redoublé par l'impératif de produire, simultanément à la version imprimée d'un livre, sa version

électronique au format ePub, avec ses métadonnées spécifiques. D'où d'ailleurs un nouveau vertige, né de la nécessité de passer de l'identification d'un livre par son simple ISBN à celle d'une « œuvre » identifiée par un code maître (l'ISTC)<sup>3</sup>, permettant de cataloguer toutes les versions (imprimées et électroniques) d'un même livre. Malgré les recommandations du SNE<sup>4</sup>, cette procédure est encore bien loin, en France comme ailleurs, de s'être normalisée. Le chantier est en effet fort complexe (et coûteux), impliquant de profonds bouleversements des pratiques éditoriales. Mais aussi interprofessionnelles, car dans cet univers en constante évolution, l'établissement de nouvelles normes en matière de métadonnées implique impérativement une collaboration étroite entre éditeurs, distributeurs, libraires et bibliothécaires.

François Gèze

<sup>1</sup> <http://www.sne.fr/sne/groupe/universitaire.html>

<sup>2</sup> <http://www.onixedit.com/fr/Accueil.aspx>

<sup>3</sup> [http://www.istc-international.org/html/international\\_agency.aspx](http://www.istc-international.org/html/international_agency.aspx)

<sup>4</sup> <http://xvm-12-234.ghst.net/dossiers-et-enjeux/numerique/normes-et-standards.html>

Éditions La Découverte

<http://www.editionsladecouverte.fr>

François Gèze est P-DG des Éditions La Découverte depuis 1982. En dehors de son activité d'éditeur, il a toujours exercé de nombreuses responsabilités dans les regroupements professionnels (président du groupe des éditeurs de sciences humaines et sociales du SNE, vice-président du CFC, membre fondateur de l'ADELCO...)

[f.geze@editionsladecouverte.com](mailto:f.geze@editionsladecouverte.com)

### Qu'est-ce que c'est ?

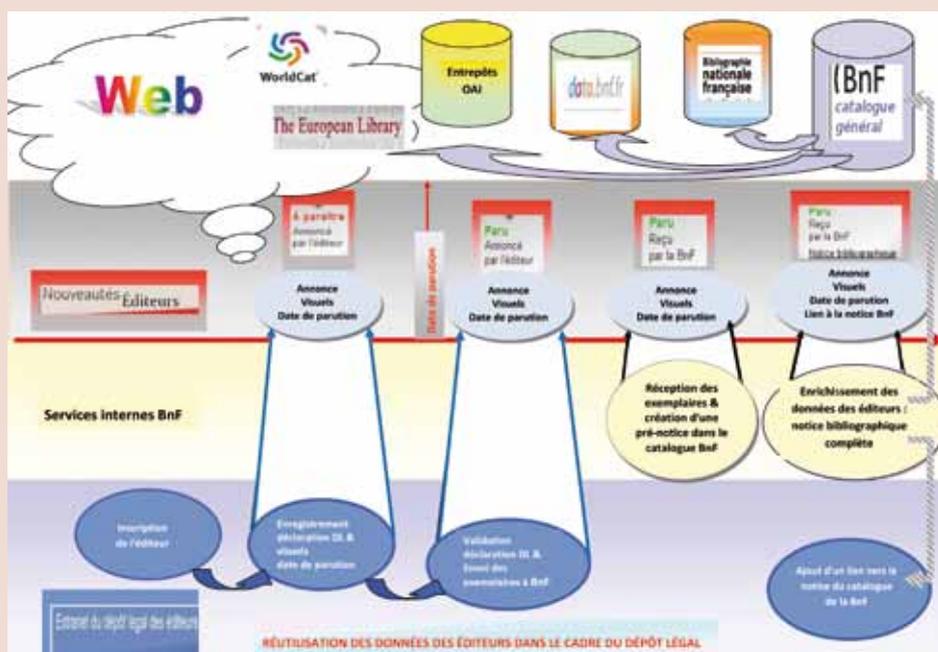
**ePub** : format de production, d'échange et de lecture pour les livres numériques, basé sur des standards web. Le contenu structuré est interprété pour une mise en forme adaptée à la lecture, quelle que soit la taille de l'écran. Au fil des versions de la norme, l'ePub permet, hormis des contenus sémantiques simples (textes, images), l'intégration et l'interprétation de contenus enrichis (dont XHTML, CSS, SVG, images, etc.). L'ePub a pour vocation de permettre la production d'un seul format numérique de publication qui alimente la chaîne de distribution jusqu'aux lecteurs finaux, autour d'un format interopérable quels que soient les moteurs de lecture et les machines qui les hébergent.

**Onix (Online Information eXchange)** : norme internationale bibliographique pour la diffusion de métadonnées enrichies. Ses principes directeurs comprennent des spécifications de contenu, d'éléments de données, d'étiquettes et de listes de codes et une DTD type XML. L'objectif est de créer un standard permettant aux éditeurs d'offrir à leurs clients une information « à valeur ajoutée » très complète (plus de 200 données possibles).

Sources : SNE, Travaux du groupe Normes & Standards, Assises du livre numérique, Salon du livre, 2012

<http://xvm-12-234.ghst.net/img/pdf/Dossiers/Normes%20et%20Standards/Guide-du-numerique-usage-editeurs.pdf>

# La réutilisation des données des éditeurs dans le cadre du dépôt légal



La réutilisation par la Bibliothèque nationale de France (BNF) des données fournies par les éditeurs est au cœur des préoccupations de l'établissement, pour faciliter la gestion du dépôt légal, accroître la visibilité des nouveautés éditoriales sur le web et établir les notices de la *Bibliographie nationale française*.

Depuis 2009, tout éditeur de livres est encouragé à s'inscrire sur le site *Extranet du dépôt légal des éditeurs*<sup>1</sup> pour y remplir ses déclarations de dépôt : date de parution prévue, titre, tirage, statut éditorial (nouveau ou nouvelle édition) sont des données obligatoires ; les informations sur les auteurs (nom, date de naissance, pseudonyme, etc.), le format du document, le façonnage, l'imprimeur, la collection éditoriale

sont optionnelles. L'éditeur peut joindre à sa déclaration électronique les fichiers contenant les visuels des première et/ou quatrième de couverture. Il est incité à faire ces saisies dans l'*Extranet* en amont de la publication du document, puisqu'il ne valide sa déclaration de dépôt que lors de l'envoi des exemplaires au département du Dépôt légal. Certaines informations de la déclaration de dépôt génèrent automatiquement une annonce descriptive, sur laquelle la BNF n'intervient pas. Depuis 2011, quand l'éditeur donne son accord, elle apparaît sous forme d'annonce accompagnée des visuels sur le site *Nouveautés Éditeurs*<sup>2</sup>. L'annonce y est publiée 3 mois avant la parution du livre et reste visible 3 mois après.

Alors que l'*Extranet du dépôt légal des éditeurs* est un site sécurisé dont l'accès est réservé aux éditeurs sur authentification, le site *Nouveautés Éditeurs* est consultable par tous ceux qui veulent s'informer sur les nouveautés de l'édition française. Ce site est mis à jour quotidiennement : au fur et à mesure que le livre « annoncé » chemine dans les chaînes de traitement de l'éditeur puis de la BNF, les qualificatifs de l'annonce évoluent de « Annoncé par l'éditeur » à « Reçu par la BNF ». Dès que les données fournies par l'éditeur ont été enrichies, livre en main, par les catalogueurs de la BNF, pour produire les notices bibliographique et d'autorité dans le Catalogue général, un lien est automatiquement généré dans les enregistrements de l'*Extranet* et de *Nouveautés Éditeurs* vers la notice bibliographique complète. Les notices de référence de livres ainsi établies sont exposées dans différents réservoirs de données, assurant à la production éditoriale française une bonne visibilité sur le web : le Catalogue général de la BNF, la *Bibliographie nationale française*, *data.bnf.fr*, les entrepôts OAI de l'établissement, etc. Elles sont largement réutilisées dans d'autres catalogues sur le web, en France, en Europe, et dans le monde.

L'*Extranet* est aussi utile aux éditeurs qu'à la BNF : c'est par là que transitent et que sont archivés les déclarations et récépissés de dépôts, les circulaires ISSN, etc. Aujourd'hui, 35 % des déclarations de dépôts de livres passent par l'*Extranet*. Pour amener plus d'éditeurs à utiliser ce service, des pistes d'amélioration sont en cours d'instruction : étendre le processus à tous les types de documents soumis au dépôt légal, récupérer les données directement dans les systèmes de gestion des éditeurs pour les dispenser de toute saisie redondante, fournir aux éditeurs les notices de référence établies par la BNF dans un format qui leur convienne et qui leur permette de les exploiter dans leurs propres circuits d'information.

*Nouveautés Éditeurs* compte aujourd'hui en moyenne 6 000 annonces en ligne, dont près d'une centaine de nouveautés par jour. Cette vitrine de l'*Extranet* s'enrichira des annonces des nouveaux types de documents, et de nouveaux types d'abonnement aux nouveautés du site sont à l'étude.

La mise en place de ces deux applications représente bien plus qu'une simple rationalisation des circuits : elle favorise les échanges entre bibliothécaires et éditeurs et préfigure l'entrée des documents numériques par dépôt légal.

Françoise Bourdon, BNF, département de l'Information bibliographique et numérique

[francoise.bourdon@bnf.fr](mailto:francoise.bourdon@bnf.fr)

Mathilde Koskas, BNF, département du Dépôt légal

[mathilde.koskas@bnf.fr](mailto:mathilde.koskas@bnf.fr)

<sup>1</sup> <http://depotlegal.BNF.fr>

<sup>2</sup> <http://nouveautes-editeurs.BNF.fr/>

# Un concert de métadonnées

**L'exploitation digitale des œuvres et des enregistrements musicaux est fortement dépendante de la façon dont les métadonnées sont renseignées et utilisées.** Celles-ci représentent donc un enjeu commun pour l'ensemble de la filière, des créateurs aux services numériques. Dans ce secteur, la gestion des métadonnées est aujourd'hui insatisfaisante, les données étant trop souvent non interopérables, erronées ou pauvres. De toute évidence, une amélioration est nécessaire, à la fois sur le plan culturel (diversité, meilleure exposition des artistes français, émergence de nouveaux talents, valorisation du patrimoine) et économique (redistribution précise et équitable des revenus, augmentation des revenus du secteur, facilitation de l'innovation et émergence d'acteurs français s'appuyant sur celle-ci). Aussi, une prise de conscience (en cours) et une mobilisation (à venir) de l'ensemble de l'écosystème musical, des ayants droit jusqu'aux acteurs du Net, est indispensable. Qu'en est-il exactement ?

## Différents types de métadonnées musicales

### Les métadonnées de propriété

Les métadonnées de propriété identifient les structures et individus ayant un droit de propriété (commerciale ou intellectuelle) sur le contenu. On les trouve sous forme nominative (nom de l'interprète) ou de codes aux normes ISO (« identifiants »)<sup>1</sup>.

Dans le secteur musical, trois identifiants<sup>2</sup> sont utilisés :

- **ISRC** (*International Standard Recording Code*) qui caractérise un enregistrement (audio et vidéo) et en indique le producteur et le pays d'enregistrement, il est généré par le producteur de l'enregistrement après attribution d'un code racine par des sociétés de producteurs (SCPP ou SPPF) ;
- **ISWC** (*International Standard Musical Work Code*) qui caractérise une œuvre (et est rattaché aux IPI et parfois aux interprètes) ;
- **IPI** (*Interested Parties Information Code*) qui caractérise un créateur ou un éditeur.

L'ISWC et l'IPI sont attribués par la Confédération internationale des sociétés d'auteurs et compositeurs (CISAC)

### Les métadonnées de gestion

Les métadonnées commerciales sont implémentées par les producteurs, les distributeurs et les sites, dans une logique qui est propre à chaque intervenant afin de faciliter leur gestion au quotidien, tant au niveau de leur propre système d'information que de celui de leurs partenaires. On retrouve ainsi les codes tarifaires, promotionnels, les territoires d'utilisation, tout ce qui peut concerner les licences d'autorisation et plus généralement toutes les codifications internes.

### Les métadonnées descriptives

Les métadonnées descriptives sont destinées à optimiser l'utilisation des fichiers par les moteurs de recherche, les sites et les applications. Elles permettent à l'utilisateur d'obtenir des résultats de recherche plus pertinents, de composer des programmes de radio/listes de lectures plus performants (sur des répertoires beaucoup plus étroits et profonds), etc. Elles peuvent être implémentées par les producteurs, les distributeurs, les moteurs de recherche, les sites, ceux qui utilisent les applications, et parfois les utilisateurs eux-mêmes.

Les métadonnées descriptives peuvent être :

- objectives (nom des différents intervenants, année de production, durée, rythme, type d'instruments utilisés, etc.) ;
- culturelles (genre et sous-genre, thématiques abordées, etc.) ;
- comportementales (« les internautes qui aiment ce fichier aiment aussi celui-là »).

### Les contenus associés

Les contenus associés peuvent être figés et liés à des contenus (traductions, photos, livrets, biographies d'artistes, interviews, discographies ou filmographies, vidéos, chansons ou scènes supplémentaires, etc.). Ils peuvent aussi prendre la forme de liens vers les sites officiels des artistes (actualités) et des tourneurs (dates de concerts), les médias officiels (articles de presse), les partenaires commerciaux (publicité, vente...), les sites sociaux (Wikipedia, Facebook), etc. Ils sont implémentés par les producteurs, les distributeurs, les moteurs de recherche, les sites, ceux qui utilisent les applications, et parfois les utilisateurs eux-mêmes.

## Une situation préoccupante

Particulièrement dans le contexte de l'exploitation de phonogrammes, les identifiants restent souvent :

- non présents : l'ISWC et l'IPI ne sont pratiquement jamais indiqués et les ISRC absents (surtout pour la création indépendante, mais également chez les majors lorsque les impératifs de promotion imposent de diffuser rapidement les enregistrements sur le Net) – à noter aussi que dans les médias traditionnels (radios hertziennes et TV) les ISRC restent sous utilisés ;
- erronés (erreur de saisie, mais aussi appropriation litigieuse) ;
- équivoques (plusieurs ISRC peuvent être en circulation suite à des repressages locaux, des compilations, des singles).

De plus, les bases ne sont pas exhaustives – en particulier pour les catalogues patrimoniaux, libres de droits – ou créées par les utilisateurs.

« **Mettre sa musique sur le net sans métadonnées, c'est comme distribuer des flyers pour un concert sans indiquer le nom de la salle** », Jean-François Bert, dans un entretien publié sur le site de l'IRMA. Phot. Juska Wendland sur Flickr (CC BY-SA 2.0)

L'absence d'identifiants peut s'avérer dramatique et engendrer :

- une rémunération des ayants droit problématique : les sociétés de gestion collective, les agrégateurs, les maisons de disques rencontrent des difficultés à rémunérer précisément les ayants droit pour les exploitations numériques. Ce qui entraîne des surcoûts de gestion, des tensions entre acteurs et, au final, des rémunérations non adéquates, voir nulles pour certains ayants droit<sup>3</sup> ;
- un frein à la bonne utilisation des autres métadonnées : un ISRC faisant défaut peut empêcher des métadonnées de différentes sources de se recouper. Il en résulte une exposition et une exploitation numérique non optimisées des enregistrements (titres non affectés aux bons artistes, doublons, artistes et enregistrements introuvables par les recherches, etc.) qui pénalisent la qualité de l'offre légale.

## De quoi avons-nous nous besoin ?

- Les identifiants doivent être présents dans tous fichiers audio ou vidéo professionnels. À défaut, ils doivent être associés aux contenus exploités dans les bases de données des acteurs de la filière.
- Les identifiants en circulation doivent être normés, univoques et fiables.
- Les métadonnées de propriété, de gestion et d'enrichissement doivent être exactes et les plus exhaustives possibles.
- Les identifiants en circulation doivent être régulièrement nettoyés.

Ainsi, les bénéfiques attendus seront de plusieurs ordres.

### Bénéfices culturels

- Lisibilité renforcée de l'offre musicale française sur les plateformes digitales françaises et internationales.
- Meilleure visibilité des répertoires fragiles (petits producteurs).
- Renforcement de l'efficacité des systèmes de gestion de droits actuels, meilleure protection contre les visées hégémoniques de certains acteurs.

### Bénéfices financiers

- Rémunération plus rapide et plus précise des ayants droit (en France et à l'international).
- Baisse importante des coûts de gestion des bases de données (producteurs, éditeurs, prestataires de services).
- Coûts d'administration des droits en baisse pour tous les acteurs.
- Revenus en hausse pour la filière : plus de métadonnées = des services plus riches et plus performants = plus grande acceptation à payer de la part de l'utilisateur.

## Développement de l'écosystème

- De nombreux services à valeur ajoutée et applications pourront être créés pour nettoyer, compléter, enrichir les métadonnées, les mettre en relation avec des contenus connexes et donner naissance à des services proposant de nouvelles expériences.
- Des appels à projets pourront être organisés sur des aspects techniques : technologie de *watermarking/fingerprinting* (empreintes), indexation automatique, web sémantique, *Open Data*, etc.
- Une filière française compétitive à l'international pourrait voir le jour autour de ces services et applications et essaïmer au-delà de nos frontières.

## Comment y parvenir ?

Des initiatives existent déjà pour optimiser les métadonnées au niveau international (DDEX<sup>4</sup>, ISRC 2, création du *Global Repertoire Database*<sup>5</sup>, base OMPI<sup>6</sup>...), mais ni celles-ci, ni le développement du web sémantique ne répondent à tous les besoins. Une action au niveau national est aujourd'hui indispensable.

Au sein de Cap Digital (pôle de compétitivité de la filière des contenus et services numériques), plusieurs structures (regroupant des ayants droit, des prestataires techniques, des acteurs du Net, etc.) se sont associées au sein d'un groupe de travail afin de faire émerger des pistes crédibles d'amélioration de la gestion des métadonnées musicales. Ce groupe, que j'ai eu le plaisir d'animer avec Elisabeth Racine, a rendu son rapport à la mission de préfiguration du Centre National de la Musique (CNM) en juin 2012. Celui-ci débouche sur deux scénarios alternatifs :

- la constitution d'une base de référence centralisée regroupant les principaux identifiants, à la gouvernance irréprochable, et agissant comme une « machine à laver » permanente mettant à disposition des identifiants certifiés pour tous les contenus musicaux ;
- ou un mécanisme contrôlé de mise à disposition des identifiants par les primo diffuseurs à des acteurs tech-



niques sélectionnés par une autorité *ad hoc*, afin que de nombreuses initiatives privées puissent éclore autour de l'amélioration de la gestion des métadonnées.

Il est encore trop tôt pour détailler des différents scénarios<sup>7</sup> et la suite qui leur sera donnée. Mais une chose est d'ores et déjà certaine : si aucune initiative européenne n'est prise, dans les années à venir, les métadonnées musicales, comme l'ensemble des métadonnées culturelles, seront de fait gérées par quelques grands groupes internationaux qui pourront alors imposer leurs règles compte tenu de leur position hégémonique dans la consommation numérique de la culture. Sommes-nous partants (ou pas) pour que la rémunération des créateurs français soit décidée dans la Silicon Valley ?

Jean-François Bert

Jean-François Bert est animateur du groupe de travail sur les métadonnées de propriété au sein du groupe Musique de Cap Digital et président de Transparency

[jfb@transparency-rights-management.com](mailto:jfb@transparency-rights-management.com)

Transparency, premier tiers de confiance spécialisé dans la gestion numérique des droits de propriété littéraire et artistique, certifie l'utilisation exacte des œuvres et enregistrements sur le Net (nature de l'utilisation, durée territoire, etc.) et renvoie aux ayants droit (société de gestion collective, producteur, éditeur, distributeur) toutes les informations nécessaires à la rémunération des créateurs, dans le format informatique souhaité par chaque client. [www.transparency-rights-management.com](http://www.transparency-rights-management.com)

<sup>1</sup> Voir aussi encadré « Des identifiants à gogo », p. 17.

<sup>2</sup> À noter qu'un nouvel identifiant ISNI est en train de voir le jour et caractérisera une « partie prenante » dans la création (ex. : créateur, producteur, mais aussi personnage de fiction, etc.).

<sup>3</sup> Pour les sociétés de gestion des droits d'auteur, les volumes de données à traiter ne cessent de croître. En 2010, le taux d'identification des œuvres numérique à la SACEM était en moyenne de 70 % sur les œuvres les plus téléchargées, les œuvres non reconnues doivent être traitées manuellement. Source : Rapport d'activité SACEM 2010, p. 29.

<sup>4</sup> <http://www.ddex.net>

<sup>5</sup> <http://www.globalrepertoiredatabase.com>

<sup>6</sup> <http://www.wipo.int>

<sup>7</sup> Au moment de la rédaction de cet article, certains points restent encore à affiner.

## Un ange passe...

L'internet des objets, cité en conclusion de l'article de Jacques Millet et Gaëlle Rivérieux comme le web de demain (cf. p. 6), rencontre l'intérêt d'un nombre croissant d'artistes.

Parmi eux, Albertine Meunier pratique l'art numérique depuis 1998 et utilise Internet comme matériau.

L'Angelino est une bouteille musicale qui détecte le passage d'un « ange » sur Twitter en repérant les messages où ce mot apparaît. À chaque apparition, la danseuse se déclenche et tourne en émettant une musique. Albertine Meunier souhaite ainsi mettre en évidence le flux de données échangées sur les réseaux sociaux et plus précisément sur Twitter.

La connexion à Internet est effectuée par l'intermédiaire d'un module Arduino qui permet de construire des objets interactifs indépendants ou de communiquer avec des logiciels.

« L'Angelino d'Albertine Meunier réussit [...] à convertir délicatement le tissage du réseau numérique en une danse d'arabesques » (Cabinet de curiosités numériques de nod-A).



L'Angelino d'Albertine Meunier a fait l'objet de plusieurs expositions, notamment dans le cadre du Carrefour des possibles au CNAM en 2009.

Phot. Pierre Métivier sur Flickr (CC BY-NC 2.0)

Site d'Albertine Meunier :

<http://www.albertinemeunier.net>

# Des identifiants normalisés à gogo

L'Internet se mue effectivement en un « web » : un réseau, un tissu. Qui plus est un web de données, dans lequel l'information se fractionne en éléments devenant autonomes et susceptibles d'être liés les uns aux autres en d'imprévisibles et potentiellement innombrables combinaisons de données<sup>1</sup>. Cela à une condition : que les données en question soient dûment identifiées par des URI (*Uniform Resource Identifier*). Une URI doit se conformer à une syntaxe particulière, mais peut utiliser tout standard d'identification existant... pourvu qu'il identifie.

Les entités bibliographiques sont de diverses natures : produits (œuvres, publications), auteurs et contributeurs, maisons d'édition, sujets (eux-mêmes de natures multiples par définition), etc. L'avènement d'un « web de données » rend nécessaire la mise au point d'identifiants standard, de préférence internationaux, pour ces différentes entités.

Les bibliothèques connaissent et font depuis longtemps usage de l'**ISBN** (*International Standard Book Number* : Numéro international normalisé de livre) et de l'**ISSN** (*International Standard Serial Number* : Numéro international normalisé de publication en série). Le premier, qui identifie une production éditoriale monographique, généralement textuelle, est décrit dans une norme internationale de l'ISO<sup>2</sup> (ISO 2108:2005<sup>3</sup>). Le second s'applique aux ressources continues et fait l'objet de la norme ISO 3297:2007. Citons dans le même ordre : l'**ISMN** (*International Standard Music Number* : Numéro international normalisé de la musique, ISO 10957:2009).

Les identifiants qui viennent d'être énumérés ont été créés avant la mise au point des modèles FRBR<sup>4</sup> et FRAD<sup>5</sup>, et se situent tous au niveau de la « manifestation ». Or ce sont à présent les œuvres et leurs expressions respectives qu'il s'agit d'identifier sur le web, de même que les personnes et/ou les collectivités qui ont contribué à leur création, leur transformation, leur mise à disposition, ou bien qui en détiennent un ou des spécimen(s). Le besoin est à la fois bibliographique, commercial et juridique (gestion des droits). D'où la création plus ou moins récente de nouveaux identifiants répondant à ces besoins.

## Œuvres et expressions

L'objectif de l'**ISTC** (*International Standard Text Code*, ISO 21047:2009) est en principe d'identifier les **œuvres textuelles** et les ayants droit associés. Dans la pratique, il peut aussi être attribué à des expressions particulières d'une même œuvre. Il en va de même, pour les **œuvres musicales**, de l'**ISWC** (*International Standard Musical Work Code*, Code international normalisé des œuvres musicales, ISO 15707:2001). Les **œuvres audiovisuelles** sont identifiées par un **ISAN** (*International Standard Audiovisual Number*, Numéro international normalisé d'œuvre audiovisuelle, ISO 15706:2002), et leurs **versions** (qui en sont des expressions selon les FRBR) par un **V-ISAN** (ISO 15706-2:2007). L'**ISRC** (*International Standard Recording Code*, Code international normalisé des enregistrements, ISO 3901:2001) s'applique aux enregistrements sonores (placés au niveau de l'expression dans les FRBR).

## Personnes et collectivités

Le tout récent **ISNI** (*International Standard Name Identifier*, Code international normalisé des noms, ISO 27729:2012) a pour application « l'identification des identités publiques des parties concernées », c'est-à-dire les **noms**, pseudonymes compris, utilisés par des personnes et/ou des collectivités impliquées dans les processus de création, de production, de distribution de contenus culturels.

## Bibliothèques

Deux identifiants spécifiques s'appliquent aux bibliothèques : l'**ISIL** (*International Standard Identifier for Libraries and Related Organizations*, Identifiant international normalisé pour les bibliothèques et les organismes apparentés, ISO 15511:2011), qui identifie les bibliothèques elles-mêmes, et l'**ISCI** (*International standard collection identifier*, ISO 27730:2012) pour leurs fonds.

Philippe Le Pape, Chargé de mission normalisation, ABES

[le-pape@abes.fr](mailto:le-pape@abes.fr)

Pour en savoir plus : Voir la page consacrée aux identifiants normalisés sur le site internet de l'ABES :

<http://www.abes.fr/Normalisation/Metadonnees-bibliographiques/Identifiants-normalises>

<sup>1</sup> Voir article de Yann Nicolas « Vers des référentiels globalisés », p. 7.

<sup>2</sup> ISO : *International Organization for Standardization*, Organisation internationale de normalisation.

<sup>3</sup> Dans les références des normes ISO le nombre précédant le signe « : » identifie la norme, le nombre suivant ce signe désignant l'édition en cours (généralement par l'année d'édition).

<sup>4</sup> FRBR : *Functional Requirements for Bibliographic records*, Spécifications fonctionnelles des notices bibliographiques (IFLA, 1998).

<sup>5</sup> FRAD : *Functional Requirements for Authority Data*, Fonctionnalités requises des données d'autorité (IFLA, 2009).



Gogo's. Phot. Tcheng Chang sur Flickr (CC BY-NC-SA 2.0)

# Métadonnées générées par les utilisateurs

## Le cas des pratiques d'indexation sociale

**L**es pratiques liées aux usages des outils du web social ont contribué à inscrire progressivement la création et l'utilisation de métadonnées dans les habitudes de nombreux internautes, de façon consciente ou non. Les indispensables opérations de description des contenus créés, partagés et diffusés sur le web – sans lesquelles les ressources concernées resteraient difficilement accessibles – poussent les utilisateurs à devenir naturellement producteurs « d'informations sur l'information ». Qu'il s'agisse de décrire le contenu d'un billet posté sur un blog, de caractériser un document sur une plateforme de diffusion (Flickr, Picasa, YouTube, Dailymotion, etc.) ou encore de partager un ensemble de signets à l'aide d'un service dédié (Delicious, Diigo, etc.), l'internaute est amené à décrire et caractériser, de fait, un ensemble de ressources documentaires. Ce processus est aujourd'hui couramment qualifié d'indexation sociale en référence aux aspects collaboratifs et communautaires de l'exercice ainsi qu'à ses finalités. Ces pratiques présentent de nombreuses particularités tant dans leurs modalités de mises en œuvre que dans la nature des résultats obtenus. L'exploitation des corpus de métadonnées ainsi produits ouvre de nombreuses perspectives, parfois fort éloignées des raisons ayant conduit à leur création. Nous proposons, dans un premier temps, d'en rappeler les principales caractéristiques, notamment par leur mise en perspective du point de vue des pratiques documentaires plus classiques. Dans un second temps, nous présenterons plusieurs contextes d'utilisation de ces métadonnées afin de mettre en exergue la richesse et la diversité des applications envisageables.

### Tags, folksonomies et pratiques documentaires

L'indexation sociale n'a pas vocation à se substituer aux pratiques professionnelles établies dans le domaine de l'information-documentation dont la rigueur reste gage d'efficacité. Elle en diffère sur de nombreux aspects. Les objectifs poursuivis ne sont pas nécessairement les mêmes. Si l'indexation a pour objet la description, au moyen d'outils documentaires<sup>1</sup>, du contenu informationnel d'une ressource, l'indexation sociale s'inscrit dans une démarche plus vaste mais moins clairement définie. L'utilisation par l'internaute de descripteurs, généralement désignés par le terme de tags, énoncés sans aucune contrainte formelle ou sémantique ouvre un champ de possibilités beaucoup plus étendu qu'une approche purement descriptive. Différents travaux de recherche ont mis en évidence le fait que, outre ce travail de description, les utilisateurs de systèmes à base de folksonomies<sup>2</sup> ont recours aux tags afin de qualifier, de commenter ou d'annoter des ressources (en utilisant des tags comme « à lire », « à faire », « intéressant », etc.). Le travail d'indexation est traditionnellement le fruit de

l'activité de praticiens de l'information-documentation formés tant à l'accomplissement de cette tâche qu'au respect des normes et des bonnes pratiques propres à leur champ disciplinaire. Dans le cas de l'indexation sociale, les corpus de métadonnées générés par les utilisateurs résultent d'une accumulation de représentations individuelles exprimées librement par des individus qui ne sont pas nécessairement sensibilisés aux bénéfices qu'apporte une démarche rigoureuse. L'usage de tels outils entraîne le déplacement de certaines pratiques documentaires vers un public novice en la matière. La très forte hétérogénéité des résultats ainsi obtenus constitue bien entendu un obstacle à partir du moment où ils sont considérés uniquement à travers le prisme de la documentation. En revanche, ils tirent toute leur richesse de l'accumulation et de la mutualisation réalisées au sein d'espaces communautaires en ligne.

L'essor de ce type de systèmes collaboratifs d'indexation de masse sur Internet s'explique en raison de :

- leur facilité d'utilisation permettant au plus large public d'appréhender leurs fonctionnalités très rapidement et sans contraintes ;
- leur capacité à créer et nourrir des échanges d'information entre utilisateurs notamment par les mécanismes de suggestion de tags ;
- l'adaptabilité du vocabulaire utilisé pour décrire et rechercher les ressources ;
- l'ensemble de solutions qu'ils proposent pour répondre aux besoins individuels d'organisation et de stockage de l'information.

Le corollaire de ces propriétés est l'absence de règles de syntaxe, l'omission du traitement des ambiguïtés langagières ainsi que l'absence de prise en compte des flexions des différents descripteurs. Ces aspects sont particulièrement problématiques dans un environnement multilingue qui amplifie considérablement leur fréquence d'apparition. Autant d'obstacles très handicapants dans le cadre d'un travail d'indexation rigoureux que l'usage d'outils spécifiques, comme les langages documentaires, permettrait d'éviter ou du moins de limiter. Doit-on pour autant considérer que les métadonnées ainsi générées ne constituent qu'un ensemble d'informations désordonné, imprécis et au final peu signifiant ? En dehors de leurs finalités premières, il nous semble que ces sources d'information offrent des potentialités d'usage multiples tant dans le domaine de la veille informationnelle et stratégique que pour l'enrichissement de bases de données partagées. Elles ouvrent également d'intéressantes perspectives sur l'évolution des outils de recherche d'information qui accordent toujours plus de place à la dimension sémantique des liens existant entre les unités informationnelles repérées sur les réseaux.

## Potentialités d'usage des métadonnées générées par les utilisateurs

Les folksonomies sont caractérisées par une très faible inertie face à la nouveauté. L'utilisation de néologismes et l'expression libre de nouveaux concepts permettent une circulation et un partage rapides de l'information et l'émergence de nouvelles acceptions, généralement liées à l'usage. Le caractère labile des métadonnées mobilisées contribue à en faire de précieux indicateurs de tendances. L'étude approfondie de l'évolution temporelle d'un corpus folksonomique, notamment des variations de fréquences d'utilisation des descripteurs et de l'apparition de cooccurrences entre descripteurs, permet de définir des processus d'identification des « signaux faibles » essentiels au travail de veille.

Ces mêmes folksonomies peuvent également participer à l'enrichissement de données structurées préexistantes. Pour des organismes partageant des ressources en ligne, le fait d'associer les usagers à la production de métadonnées permet d'envisager une démarche de « redocumentarisation » particulièrement intéressante. Les fonds partagés et, par extension, les bases de données permettant de les gérer peuvent ainsi être enrichis d'une couche supplémentaire d'information ayant pour objet, au moyen d'une personnalisation accrue, l'augmentation des interactions entre usagers et contenus numériques.

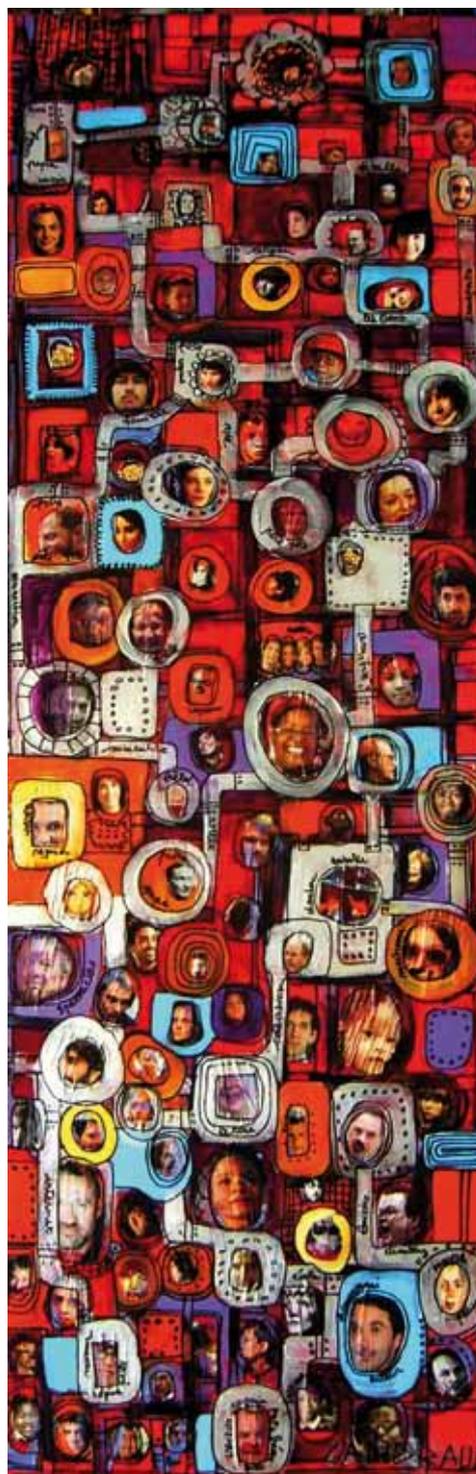
Enfin, si elles sont parfois présentées comme deux approches antagonistes, la conception d'un « web social » privilégiant les applications ouvertes, communautaires et partagées, et celle d'un « web sémantique » organisé, rigoureux et formalisé, ne semblent plus aujourd'hui s'exclure totalement. Elles peuvent même nourrir d'intéressants points de convergence. Les futurs outils de recherche et d'accès à l'information, en particulier pour leurs fonctionnalités de personnalisation et de filtrage des résultats proposés, auront tout intérêt à tirer le meilleur parti à la fois des schémas de métadonnées formalisés retenus comme éléments constitutifs d'une architecture raisonnée de l'information mais également de la richesse et de la diversité apportées par les métadonnées générées par les utilisateurs.

Fabrice Pirolli

<sup>1</sup> Lexiques, listes d'autorités ou thésaurus.

<sup>2</sup> Néologisme issu de la contraction des termes « folks » et « taxonomies » désignant le principe d'une indexation libre, de masse et partagée.

Fabrice Pirolli est maître de conférences en sciences de l'information et de la communication, Université de Bourgogne, et chercheur au laboratoire CIMEOS EA 4177  
[fabrice.pirolli@iut-dijon.u-bourgogne.fr](mailto:fabrice.pirolli@iut-dijon.u-bourgogne.fr)



### Quelques indications bibliographiques

- Casemajor Loustau, Nathalie, « La contribution triviale des amateurs sur le Web : quelle efficacité documentaire ? », *Études de communication*, n° 36, 2011, p. 39-52.
- Filippi, Dominique, « Délicieuses folksonomies : comment gérer les signets de la bibliothèque dans Delicious ? », *Le web 2.0 en bibliothèque. Quels services ? Quels usages ?*, coll. Bibliothèques, éditions du Cercle de la librairie, 2009, p. 85-113.
- Francis, Élie, Quesnel, Odile, « Indexation collaborative et folksonomies », *Documentaliste-Sciences de l'Information*, n° 1/2007 (Vol. 44), ADBS, p. 58-63.
- Pirolli, Fabrice « Pratiques d'indexation sociale et démarches de veille informationnelle », *Études de communication*, n° 36, 2011, p. 53-66.

# Le fonds des Nations unies de la bibliothèque Cujas

## Un exemple de valorisation

**Élaborer une politique de valorisation nécessite la connaissance des fonds, l'établissement de conditions minimales de conservation, la garantie de leur accessibilité.** Elle n'a de sens que par rapport à ses publics, qu'il faut, en conséquence, bien connaître dans leur diversité et bien définir lors de la mise en œuvre d'actions particulières. Enfin, tout travail de ce type doit relever d'une politique globale de l'établissement concerné et s'intégrer dans le cadre de programmes bien délimités. Depuis quatre années, c'est cette stratégie qui a été appliquée pour valoriser le **fonds des Nations unies**, détenu par la bibliothèque interuniversitaire Cujas depuis 1949 et enrichi jusqu'en 2010.

### Une continuité historique

Le 23 mars 1948, le comité des publications de l'ONU, s'appuyant sur l'ancien système créé en 1938 par la Société des Nations (SDN), établissait un réseau de destinataires bientôt soumis à deux textes cadres : les « principes relatifs aux bibliothèques dépositaires des documents et publications de l'organisation des Nations unies », en date du 20 décembre 1955, et les « instructions à l'intention des bibliothèques dépositaires », du 27 mars 1964. Ces textes furent révisés régulièrement jusqu'aux versions actuellement en vigueur, respectivement des 18 août et 5 mars 1995. Le but de ce réseau est que « le public puisse avoir accès dans le monde entier aux documents et publications de l'Organisation des Nations unies ». Les bibliothèques nationales et parlementaires des pays reçoivent gratuitement la documentation, tandis que les autres dépositaires acquittent une contribution annuelle « au titre des dépenses du fonctionnement du système ». Au sortir de la Deuxième Guerre mondiale, Cujas entre donc dans ce réseau avec trois autres bibliothèques : l'actuel SCD de l'université Bordeaux 4, la BNU de Strasbourg et Sciences-Po

Paris. Ce dépôt fut négocié par Mme Suzanne Bastid (1906-1995)<sup>1</sup>, professeur de droit international à la Faculté de droit de Paris, membre du tribunal administratif des Nations unies en 1949, tribunal qu'elle préside ensuite de 1953 à 1963. Le fait d'être dépositaire conforte un positionnement ancien de la bibliothèque Cujas qui procédait déjà, en 1921, à des échanges avec le Bureau international du travail (BIT) et la SDN ou, en 1923, avec la Cour permanente de justice de la Haye, à l'initiative de Ferdinand Larnaude, doyen de la Faculté de droit de Paris, et d'E. Bouvy, bibliothécaire en chef. Dans les années 1940-1950, la bibliothèque a enrichi ses collections des dépôts de l'OCDE, de la Banque mondiale, du FMI, de l'OTAN, de l'Union de l'Europe occidentale (UEO), de l'Organisation des Nations unies pour l'alimentation et l'agriculture (FAO), de l'Organisation mondiale pour la propriété intellectuelle (OMPI), etc., tandis qu'en 1963, elle devient également Centre de documentation européenne.

### Un service spécifique

Un **service dédié** (dès 1956 au moins) fut créé et, durant plusieurs dizaines d'années, compta de deux à cinq personnes. Tout apparentait le service des publications internationales à une petite bibliothèque dans la bibliothèque. Réception, bulletinage, inventaire, catalogage, traitement des factures et des réclamations, mise à disposition, renseignement au lecteur lui incombaient. Une partie des collections se trouvait en libre accès dans une petite salle ouverte en 1988 ; une autre part était conservée dans une réserve propre ; le reste observait le circuit normal des périodiques et monographies. En 1981, on comptabilise 1 100 demandes pour 3 500 documents de la part de 380 lecteurs pour les documents de travail des Nations unies<sup>2</sup>. Les années 1990 sont un tournant. L'essor d'Internet et de ressources nombreuses mises en ligne par les Nations unies elles-mêmes marquent une désaffection du

lectorat et un relatif déclin du service. Celui-ci devient un corps à part dans l'établissement, submergé par des centaines de dons (281 titres en 2009) dont plusieurs sans réel intérêt pour les disciplines de la bibliothèque.

### État des lieux et valorisation

En 2008, une étude fut lancée pour réintroduire le service dans le circuit normal de la bibliothèque. Un audit réalisé en interne a permis de distinguer deux axes de travail : former/informer sur les ressources internet et mettre en valeur le caractère patrimonial de la collection. Au préalable, il a fallu adapter les dons reçus aux usagers de la bibliothèque et, en priorité, aux doctorants et chercheurs en droit international ou en sciences politiques. Une étude approfondie, organisme par organisme, a été enclenchée pour délimiter un nouveau cadre documentaire, hors duquel les publications reçues ne seraient plus conservées. L'année 2009 vit ainsi la suppression du dépôt de quelques organismes satellites des Nations unies, et, dans un second temps, de celui des documents officiels, s'inscrivant ainsi dans les orientations de l'ONU pour une politique d'offre en ligne de plus en plus étendue *via* son site Unbisnet<sup>3</sup>. De nouvelles acquisitions permirent ensuite d'étayer des thématiques précises restées lacunaires. Enfin, les amples ressources électroniques auxquelles la bibliothèque est abonnée consolident encore l'offre courante. Le libre accès a été réorganisé, passant de la petite salle dédiée peu fréquentée à des rayonnages spécifiques au sein de la salle de lecture générale, rendant plus visibles les ouvrages de référence et les instruments de recherche. En parallèle, un *Guide des documents de l'ONU* (mai 2011), en accès libre sur le portail de la bibliothèque, régulièrement présenté aux lecteurs mais aussi en interne aux collègues de l'établissement, a été rédigé en tenant compte des changements rapides des

technologies de l'information et de la communication ainsi que des nouveaux moyens de diffusion de l'information.

## L'inventaire via Calames

Il restait à valoriser les cartons d'archives, quelque peu tombés dans l'oubli depuis une dizaine d'années. Ces archives sont constituées des documents mimeographés (ronéotypés), à savoir les documents de travail des grands organes de l'ONU : Assemblée générale, Conseil de sécurité, Conseil de tutelle, Conseil économique et social, Secrétariat et quantité d'autres fonds ou programmes onusiens : Programme des Nations unies pour le développement (PNUD), Conférence des Nations unies sur le commerce et le développement (CNUCED), FAO, Agence internationale sur l'énergie atomique (AIEA), etc., stockés dans 3 788 boîtes, sur près de 400 mètres linéaires. Suite à un audit de la mission archives du ministère de l'Enseignement supérieur et de la Recherche, le projet de traiter ce fonds dans Calames a été lancé. Après un dédoublement réalisé à l'été 2010 (40 ml pilonnés soit 429 cartons) eut lieu, durant 4 mois, le récolement à partir des indications portées sur les cartons, héritage précieux du travail réalisé par les bibliothécaires successifs qui reportaient avec précision sur chaque boîte l'organisme émetteur, les dates butoirs, parfois quelques mots-clés et, surtout, les cotes ONU correspondantes, dans le respect de son **système de classification unique**. Cette nomenclature donne un langage commun avec nos lecteurs spécialistes de ce domaine. Ce récolement a permis ensuite de construire l'inventaire par organismes émetteurs et, enfin, d'en réaliser la saisie sur l'outil Calames. Cette avancée est extrêmement positive à plus d'un titre : le fonds sort de son anonymat ; la grande précision dans le descriptif des notices et l'ergonomie très souple de la base Calames permettront des enrichissements ultérieurs en cas de dépouillement plus fin des cartons ; enfin, il est possible de faire des



La salle de réunion du Conseil de sécurité des Nations unies à New York.  
Phot. Bernd Untiedt sur Wikimedia Commons (CC BY-SA 3.0)

recherches par mots-clés, par feuilletage de l'inventaire et, aussi par cote ONU, ce qui était le véritable enjeu de ce catalogue en ligne. Cette valorisation s'est accompagnée d'une meilleure conservation par un reconditionnement en réserve en boîtes de carton neutre. Une enquête, réalisée en mai 2010 et relayée par le *Bulletin de l'école doctorale de droit international et européen* (n° 77, juin 2010) de l'université Paris 1, a montré un besoin d'informations des étudiants sur l'existence de ce fonds. Notre projet est d'intensifier les liens avec des centres de recherche ou des laboratoires susceptibles d'être intéressés par notre documentation et de l'exploiter scientifiquement. Des demandes récentes de chercheurs démontrent déjà tout son potentiel. Ainsi, il y a peu, les cartons du Haut Commissariat des Nations unies pour les réfugiés (UNHCR), sur la période 1949-1993, ont été dépouillés par un enseignant de Sciences-Po Bordeaux, en vue de réaliser une base de données sur ce thème. À Paris, un doctorant a

pu consulter les travaux du Conseil de sécurité et de la CNUCED, entre 1946 et 1980, à l'appui de sa thèse portant sur le Brésil et le Mexique dans le multilatéralisme onusien depuis 1945. Sans avoir la prétention de se substituer aux fonds d'archives de Genève ou de New York, il est bien certain que la bibliothèque Cujas offre actuellement, dans le paysage documentaire français, le fonds sur les Nations unies le plus complet, et maintenant bien inventorié.

Catherine Désos-Warnier

<sup>1</sup> Note à Gloria M. Leo, chef de section des acquisitions à New York, 20/02/1981 [archives du service]. Mme Bastid est la fille de Jules Basdevant et l'épouse du ministre Paul Bastid.

<sup>2</sup> A. H. Weber, conservateur en chef à la BIU Cujas, à Cornel Metternich, directeur de l'*United Nations Information Center Paris*, 23/09/1982 [archives du service].

<sup>3</sup> *UN Bibliographic Information System* : <http://unbisnet.un.org/>

Bibliothèque interuniversitaire Cujas

<http://biu-cujas.univ-paris1.fr>

Catherine Désos-Warnier est adjointe au chef de département des périodiques

[Catherine.Desos-Warnier@univ-paris1.fr](mailto:Catherine.Desos-Warnier@univ-paris1.fr)



# ADBU / AURA : on fusionne !

## Les premiers jalons

Lors du congrès de Vannes en septembre 2011, l'ADBU a pris à son compte, à l'initiative de l'AURA et en coopération avec Couperin, la réflexion sur la réorganisation du paysage associatif professionnel de l'information scientifique et technique (IST). Missions et objectifs partiellement redondants, positionnement de l'AURA à redéfinir en raison des évolutions de l'ABES ces dernières années, modification du paysage national avec la loi relative aux libertés et aux responsabilités des universités (LRU), difficultés à recruter pour les conseils d'administration (CA) dans les mêmes viviers de professionnels déjà fortement sollicités par ailleurs, nécessaire intégration des organismes de recherche, urgence toujours repoussée à dégager les moyens d'une véritable professionnalisation du travail associatif : les sujets ne manquent pas pour alimenter la discussion.

## Trois scénarios

Les CA de l'ADBU et de l'AURA ont travaillé, durant l'automne 2011 et l'hiver 2012, sur les hypothèses de rapprochement des deux associations et rencontré à plusieurs reprises les différents acteurs du projet : ABES, Couperin, MISTRD notamment. Les travaux ont abouti à un document de synthèse diffusé auprès des membres des deux associations pour alimenter le débat démocratique. Le document exposait trois scénarios d'évolution possible, présentés par ordre d'intérêt croissant : statu quo ; développer les projets en inter-association ; fusionner l'AURA et l'ADBU. Un calendrier d'action en plusieurs phases complétait le plan d'action proposé.

La première étape a eu lieu le 6 avril 2012 sous la forme d'assemblées générales extraordinaires des deux associations, tenues le même jour successivement. Les deux assemblées ont répondu « oui » à de très larges majorités à la proposition des CA (*Vous prononcez-vous en faveur de l'intégration de l'AURA et de ses missions dans l'ADBU ?*) et lancé les travaux d'adaptation des statuts de l'ADBU.

## Cinq chantiers

Les débats en assemblées ont permis de décliner cinq axes de travail, c'est-à-dire cinq chantiers concrets pour aboutir à une révision complète des statuts de l'ADBU ou à une proposition de statuts pour une nouvelle association :

- **la redéfinition des missions de l'association** : intégration des missions de l'AURA, élargissement plus marqué vers les enjeux de la recherche, sans oublier ceux de l'enseignement supérieur ni les problématiques métiers et compétences ;
- **l'élargissement du périmètre** en ouvrant l'association à tous les membres de l'AURA et aux établissements publics à caractère scientifique et technologique (EPST) pour couvrir l'ensemble du champ de l'IST en France. Si l'AURA est une association d'établissements, alors que l'ADBU est une association de personnels de direction de bibliothèques d'uni-

versités, le périmètre du type de structures concernées est également différent : l'AURA intègre par nature les utilisateurs des produits de l'ABES, donc également des collectivités territoriales, des organismes de recherche, des établissements d'enseignement supérieur privés, qui ne doivent pas se retrouver lésés par ce rapprochement ;

- **le type d'adhésion à l'ADBU nouvelle** : individuelle (et de quel type de personnel), de service, d'établissement en prenant en compte les conséquences des diverses hypothèses (représentativité, modes d'élection au CA, montant des cotisations, etc.) ;
- **les modes d'action de l'association**, par le biais notamment d'une cinquième commission « réseau documentaire », d'une meilleure communication en direction des membres et, à terme, du renforcement des moyens humains par le recrutement d'un permanent ;
- **la structuration du travail entre associations**, en proposant d'organiser congrès et séminaires communs, commissions mixtes et groupes de travail *ad hoc* avec les autres associations du champ de l'IST, de la documentation, des archives, à l'instar de ce qui se fait couramment dans d'autres pays.

Le collectif chargé de préciser les hypothèses de travail et préparer un calendrier et une nouvelle mouture des statuts doit rendre ses conclusions et propositions d'ici fin juin 2012. La dernière étape consistera en une nouvelle assemblée générale extraordinaire, au cours de laquelle les nouveaux statuts de l'association seront proposés aux membres pour adoption et où l'AURA votera sa dissolution.

## Un but

Poursuivi depuis quelques années en dépit des difficultés de tous ordres et réitéré comme une ardente obligation dans le rapport moral 2011 de l'ADBU, l'objectif final est en fin de compte de **passer d'une association de professionnels à une association professionnelle**. L'intégration de l'AURA à l'ADBU marquera une étape importante dans ce processus évolutif : elle permettra en effet d'augmenter la représentativité de l'association, en renforçant son rôle de partenaire des tutelles et des opérateurs nationaux de l'IST ainsi que des autres associations professionnelles de l'enseignement supérieur et de la recherche. Une ADBU élargie et refondée gagnera en légitimité pour traiter des questions stratégiques de définition et d'organisation du réseau documentaire national et faire en sorte que les bibliothèques de l'enseignement supérieur et de la recherche y jouent pleinement leur rôle.

Marc Martinez, secrétaire général de l'ADBU

 [marc.martinez@ens-lyon.fr](mailto:marc.martinez@ens-lyon.fr)

Françoise Truffert, secrétaire générale de l'AURA

 [francoise.truffert@univ-valenciennes.fr](mailto:francoise.truffert@univ-valenciennes.fr)

# L'ABES au World Wide Web 2012

Rendez-vous annuel de tous les acteurs du web depuis sa première édition en 1994, W3C souhaite pérenniser les échanges entre les spécialistes et les utilisateurs afin d'anticiper l'avenir d'Internet et ses conséquences sur la société. La manifestation, organisée en 2012 par l'université de Lyon (16-20 avril), a regroupé environ 2 300 personnes, représentant près de 60 nationalités.

Programmés dans le cadre de différents dispositifs (séances plénières animées par des personnalités du web, *workshops* sur de nouveaux outils ou des problématiques, tutoriaux pour s'initier à de nouvelles technologies, conférences scientifiques plus orientées vers la recherche), les principaux thèmes abordés ont été : les réseaux sociaux, les nouvelles technologies, l'accessibilité, le respect de la vie privée, l'*Open Data* et le web sémantique. Le choix des contributions reste assez sélectif car 12 % seulement des propositions sont retenues.

Si cet événement est une bonne approche des évolutions du web, de ses impacts technologiques, humains et sociologiques, on peut toutefois regretter que certains thèmes abordés étaient difficilement accessibles aux non-initiés et que leur application concrète dans le cadre d'un établissement reste peu flagrante.

Le prochain congrès W3C se tiendra du 13 au 17 mai 2013 à Rio de Janeiro.



Jean-Laurent Alexis, ABES  
[jean-laurent@abes.fr](mailto:jean-laurent@abes.fr)



Photo : Clément Renaud sur Flickr (CC BY-NC-SA 2.0)

## Des chiffres & des lettres



### Système universitaire de documentation

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de recherches par l'interface web : 3 769 950
- Nombre de demandes de prêts entre bibliothèques : 42 930
- Nombre de demandes de prêts entre bibliothèques satisfaites : 38 418 (soit 89,5 %)

État de la base au 31 mai 2012

- Nombre de notices bibliographiques localisées : 10 573 599
- Nombre de notices d'autorité : 2 517 853
- Nombre de localisations : 34 735 032

<http://www.sudoc.abes.fr>



### Moteur de recherche des thèses de doctorat françaises

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de visites : 367 604
- Pages consultées : 1 321 338

Total des thèses signalées au 05 juin 2012 : 76 047 dont 64 816 en préparation (Step) et 11 231 soutenues (Star)

<http://www.theses.fr>



### Catalogue en ligne des archives et manuscrits de l'enseignement supérieur

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de visites : 33 379

Total des composants publiés au 31/05/2012 : 308 623

<http://www.calames.abes.fr>



### Numérisation des manuscrits de l'enseignement supérieur

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de visites : 1 606

Total des corpus numérisés au 31 mai 2012 : 191

<http://www.numes.fr>



### Signets des universités

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de visites : 10 681
- Pages consultées : 35 422

Total des signets créés au 31 mai 2012 : 2 776

<http://www.signets-universites.fr>



### Guichet d'assistance de l'ABES

Du 1<sup>er</sup> mars au 31 mai 2012

- Nombre de demandes d'assistance : 1 988 (95,7 % traitées)

<https://stp.abes.fr>



### Consultation des blogs de l'ABES

Nombre de visites du 1<sup>er</sup> mars au 31 mai 2012

- **Fil ABES** : 10 291 <http://fil.abes.fr>
- **Formation** : 2 021 <http://abesformation.wordpress.com>
- **Punktokomo** : 2 124 <http://punktokomo.abes.fr>
- **rda@abes** : 4 673 <http://rda.abes.fr>

# Agenda

## Août

 Du 6 au 9

### ISKO 2012 : Mysore (Inde)

La 12<sup>e</sup> édition du congrès international de l'ISKO (International Society for Knowledge Organization) sera consacrée à l'indexation, l'organisation des liens et la contextualisation de l'organisation des connaissances.

 <http://www.isibang.ac.in/~isko/>

 Du 11 au 16

### 78<sup>e</sup> Congrès de l'IFLA Helsinki (Finlande)

« Les bibliothèques aujourd'hui ! Inspirantes, surprenantes, rendant plus fort ». La numérisation et les médias sociaux ont bouleversé le modèle traditionnel de la bibliothèque. Cette situation ouvre un énorme défi. Le congrès de l'IFLA encouragera les bibliothèques à collaborer et à travailler ensemble pour trouver de nouveaux partenaires hors du monde des bibliothèques.

 <http://conference.ifla.org/ifla78>

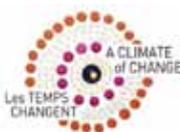


 Du 20 au 24

### ICA 2012 : International Council on Archives Congress Brisbane (Australie)

Les temps changent : le programme de ce congrès est construit autour des défis auxquels sont confrontés dorénavant les archivistes et documentalistes pour préserver, gérer et diffuser leurs collections.

 <http://www.ica2012.com/index.php>



## Octobre

 Du 1<sup>er</sup> au 3

### EMTACL12 : Emerging Technologies in Academic Libraries

#### Trondheim (Norvège)

Congrès international consacré aux technologies émergentes dans les bibliothèques universitaires.

 <http://emtacl.com/>

Photo : Matthew Rees  
sur Flickr



 Du 5 au 7

### ADBU 2012 : 42<sup>e</sup> Congrès Toulouse - Centre de Congrès P. Baudis

Pour cette cuvée 2012, la journée d'étude du vendredi 6 septembre, ouverte à l'ensemble des professionnels des bibliothèques et de la documentation, sera plus particulièrement consacrée aux métiers (et à leurs évolutions) dans les bibliothèques universitaires.

Comme à l'accoutumée, la première journée du congrès sera consacrée aux travaux internes de l'association (assemblée générale et rapport des commissions) et le congrès s'achèvera le vendredi 7 septembre par des interventions de représentants du ministère de l'Enseignement supérieur et de la Recherche, de la Conférence des présidents d'université et de l'Inspection générale des bibliothèques.

Le salon professionnel rassemblera une quarante de fournisseurs et partenaires des bibliothèques universitaires.

Informations pratiques (programme, bulletin d'inscription, accès) sur le site de l'ADBU.

 <http://adbu.fr/toulouse2012/>



 Du 12 au 14

### ETD 2012 : 15<sup>e</sup> Symposium international sur les thèses et les mémoires électroniques Lima (Pérou)

Comment « intégrer les aspects culturels dans la création et l'échange des connaissances » ?, telles seront les questions abordées lors ce symposium organisé par NTLTD (Networked Digital Library of Theses and Dissertation).

 <http://www.etd2012.edu.pe/en/index.asp>

 Du 24 au 26

### M-Libraries Conference 2012 Milton Keynes (Angleterre)

Cette 4<sup>e</sup> édition de M-Libraries traitera des impacts des technologies mobiles sur la vie quotidienne et sur le fonctionnement des bibliothèques.

 <http://m-libraries.org/>

