



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique : oubli dans l'âme ou supplément d'âme ?

Chers adhérents,

Qui a peur du numérique ?

Cette question éveille chez nous comme un sentiment de déjà entendu...

La révolution numérique est comme une onde de choc qui transforme nos modes de vie et conditionne notre avenir. Sans doute sommes-nous à ce jour trop bercés ou secoués par cette vague pour en mesurer toute l'ampleur.

C'est ce qu'a tenté de faire le magazine *Books* dans son dossier au titre provocateur : *Internet rend-il encore plus bête ?* Derrière cette question sont interrogées autant la peur que la fascination pour Internet.



L'écriture produit « l'oubli dans l'âme »

On peut lire, dans l'article de Joaquin Rodriguez¹, que ce n'est pas la première fois dans la brève histoire de la connaissance que l'homme a été confronté à ce genre d'interrogation. Bien avant l'avènement d'Internet, l'homme s'était déjà affolé sur les conséquences de ce nouvel outil mis à la disposition de la propagation de la pensée : l'écriture.

Il est intéressant de se transporter dans un des textes fondateurs de la philosophie, *Phèdre* de Platon, pour y déceler de nombreuses analogies avec le débat qui nous préoccupe ici. Il suffirait d'y remplacer le mot *écriture* par le mot *numérique* pour que cette discussion nous semble étrangement contemporaine.

¹ Je renvoie à son article publié dans le N° 7 du magazine *Books* (juillet-août 2009) : *L'Édition 2.0 : Socrate dans l'hyperespace*, et plus généralement au dossier consacré par ce même magazine à : *Internet rend-il encore plus bête ?*



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Rappelons brièvement que pour la pensée socratique le dialogue entre deux interlocuteurs permettait à l'individu de se forger un jugement. Or l'écriture, par son usage, allait bouleverser cet ordre des savoirs. Socrate déplorait ainsi « *le défigement de la connaissance et de la culture au contact de l'écriture* » selon Joaquin Rodriguez.

Dans *Phèdre*, Socrate compare l'écriture à la peinture, dont les personnages « *se tiennent debout comme s'ils étaient vivants* », mais, dès qu'on les interroge, « *ils restent figés dans une pose solennelle et gardent le silence.* »

Plus loin, l'écriture y est présentée comme un art qui « *produira l'oubli dans l'âme de ceux qui l'auront appris, parce qu'ils cesseront d'exercer leur mémoire (...). Quant au savoir, c'en est la semblance que tu procures à tes disciples, non la réalité. Lors donc que, grâce à toi, ils auront entendu parler de beaucoup de choses, sans avoir reçu d'enseignement, ils sembleront savoir beaucoup de science, alors que, dans la plupart des cas, ils n'auront aucune science ; de plus, ils seront insupportables dans leur commerce, parce qu'ils seront devenus des semblants de savants, au lieu d'être des savants.* »

L'écriture apparaît dans ces textes comme source d'oubli, féconde en chimères : elle porte atteinte à la mémoire, elle va à l'encontre de la transmission du savoir, quant elle ne va pas à l'encontre même de l'idée de savoir. A l'inverse de ce qui pourrait la caractériser aujourd'hui, alors même que ces reproches s'appliqueraient à tout ce qui relève du cyberspace.

On peut, avec Joaquin Rodriguez, imaginer un « *Phèdre, tolérant, persuadé de l'innocence du discours de Socrate* » qui « *le laisserait discourir, écrivant de son côté sur une tablette ou un papyrus à quelque ami de son âge pour lui raconter l'extravagance et l'ennui de ce bonhomme offusqué par une invention qui lui est étrangère...* ». Normal. Phèdre était un natif de l'écrit, de la même manière que l'on peut parler aujourd'hui de natifs numériques.

Le livre tuera la cathédrale

Dans le même dossier de *Books (Internet rend-il encore plus bête ?)*, Umberto Eco convoque une autre figure : celle du personnage du prêtre Claude Frollo qui, un livre imprimé posé sur la table devant lui, chuchote : « *Ceci tuera cela* ». Rappelons le décor littéraire et architectural de l'oeuvre dont est tiré le personnage : il s'agit de Notre-Dame de Paris de Victor Hugo.

Le livre tuera la cathédrale. Umberto Eco nous dit que la cathédrale médiévale « *était la seule chose qui pût faire connaître aux masses les histoires de la Bible, la vie du Christ et des saints, les principes moraux, voire les épisodes de l'histoire nationale ou les notions les plus élémentaires de géographie et de sciences naturelles* ».

L'auteur compare d'ailleurs la cathédrale médiévale à une « *sorte de programme de télévision permanent et immuable, censé apprendre aux gens tout ce qui était indispensable à leur vie quotidienne et à leur salut éternel* ». Si l'on retire à ce programme l'aspect religieux en y intégrant le principe de la laïcité, il y a là l'amorce de cet autre concept de cathédrale globale : la bibliothèque, qui s'est voulu un temps mémoire du monde, puis a cherché à être miroir du monde.

Si les cathédrales étaient perçues comme un grand livre illustré par leurs contemporains, elles nous impressionnent aujourd'hui par l'épaisseur de leur mystère. Si l'écriture et l'imprimerie ont sonné la fin d'une époque, rien n'empêche d'imaginer qu'il peut se passer la même chose avec Internet et qu'un jour peut-être on regardera les bibliothèques avec un sentiment d'étrange familiarité. Même si les passages d'une époque à une autre ne se font jamais de façon aussi tranchée.



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Partant toujours de l'idée que nous baignons aujourd'hui dans le liquide amniotique de l'ère numérique, nous pouvons aussi nous amuser à descendre le temps plutôt qu'à le remonter. Cela nous permettra de nous installer dans une vision d'avenir où, selon Joaquin Rodriguez, « *Nos pauvres cerveaux analogiques sont peut-être en train de se convertir en cerveaux numériques, bien qu'aucun neurolinguiste ne s'aventure à nous dire ce que nous gagnons ou perdons au change.* »

Le disque est mort, mais l'ombre du disque plane encore sur nous

Certains décors de cinéma nous donnent parfois l'impression que tout y est d'époque : les voitures, les maisons, les affiches, les meubles, les habits, les coiffures... Chaque objet rappelle dans les moindres détails le moment où il doit se situer dans l'histoire. Or si nous devons représenter notre époque actuelle, nous verrions qu'elle est entrecassée d'éléments de toute sorte : ceux du passé, ceux qui n'auront pas de postérité et ceux qui marqueront l'époque comme une effigie.

De la même manière, le disque restera encore entrecassé dans le 21^e siècle, même si, à l'évidence, il ne représentera plus le modèle dominant de consommation (et d'appropriation) de la musique.

Les derniers clivages entre ressources numériques et supports physiques sont en train de tomber : nous pouvons désormais télécharger de la musique dans des formats équivalents voire supérieurs au disque compact, la documentation figée qui accompagnait dans les livrets (lorsqu'ils existaient) nos chers disques ne résistera pas longtemps aux assauts des ressources documentaires interactives telles qu'on peut les trouver en ligne. Surtout l'usage nomade sera sans doute le comportement d'écoute musicale le plus fréquent.

Probablement que la logique de flux va s'imposer avec le temps. Quel sens cela aurait de résister à la possibilité de s'abonner à des millions de références musicales (Musicme) ou plus simplement encore à l'écoute de bases de données musicales (Jiwa, Deezer...) ? A moins que ce ne soit finalement le principe de la licence globale qui finisse par s'imposer. En tout cas la culture numérique est en train de prendre ses aises avec notre époque et pour longtemps.

Quelques irréductibles fétichistes, parmi lesquels je figure bien volontiers, utiliseront le 33 tours ou le disque compact comme un totem ou le symbole d'un moment d'écoute ou de méditation particulier, de même qu'ils s'abandonneront à une approche sensuelle de l'objet.

Rappelons-nous qu'au tournant du 20^e siècle, les technologies d'enregistrement sonore ont fait surgir un autre rapport à la musique et au son. Certes, on peut toujours dire que les enregistrements « *produiront l'oubli dans l'âme de ceux qui les auront écoutés, parce qu'ils cesseront d'exercer leur mémoire* », mais surtout la fixation sonore nous a permis de rendre le passé présent, le lointain proche, de nous faire un vrai cadeau de l'espace/temps précisément.

L'enregistrement a ouvert une nouvelle époque car il a permis l'écoute différée de la musique, une différence dans l'écoute également, si on compare cette manière de toucher à la musique au concert.

Que cette façon de différer l'écoute de la musique se joue dans un environnement où l'on fait appel à des supports physiques ou à des accès numériques en deviendrait presque accessoire.

Ce qu'il importe de préserver, c'est de garder à la musique toute la disponibilité de l'esprit, de l'oreille, toute notre capacité à suspendre l'action ou à se fondre dans l'écoute, de se laisser gagner tant par l'énergie, l'émotion musicale, que par ses aspects les plus méditatifs.

Musique numérique : oubli dans l'âme ou supplément d'âme ?

Qui sait, peut-être que les technologies numériques sont en train de réinventer les nouveaux dialogues socratiques ? Une nouvelle façon de former son jugement, son goût ou son attachement pour la musique. Nos interlocuteurs dans le monde numérique ont beau être distants, ils sont loin d'être « figés »



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique : oubli dans l'âme ou supplément d'âme ?

dans une pose solennelle en gardant le silence », bien au contraire ils s'exposent virtuellement, sortent d'eux-mêmes pour construire un savoir à force de dialogue.

Nous percevons déjà comment ces nouvelles technologies pourraient redéfinir les règles de la création, de l'échange, des droits d'auteurs, de la collecte et de la validation des contenus. Cela fait aussi surgir des convoitises, apparaître des mégalomanies, voire des tentations totalitaires. Faut-il ranger les projets de Google visant à un monde du savoir globalisé au rang de ces dernières ? Il me semble qu'il s'agirait là d'un jugement expéditif. Donnons du temps au temps pour faire la part des choses.

La question « Qui a peur du numérique ? » est donc, vous l'aurez compris, une fausse question.



Il me semble que, faute de recul précisément, l'enjeu pour les bibliothèques se concentre dans ce que Joaquin Rodriguez définit comme étant le rôle d'adultes. Il s'agirait « d'essayer de réguler et d'organiser l'usage des moyens numériques de création et de distribution, en le complétant par l'enseignement et la pratique de la lecture traditionnelle, afin que les cerveaux de nos enfants ne soient pas unipolaires, mais bitextuels, capables d'évoluer avec la même aisance sur les deux supports, et de jouir des avantages de chacun. Le rôle des professionnels de l'édition et de la culture écrite est de comprendre avec la plus grande précision et profondeur possible la nature et l'ampleur du changement en cours. »

Si d'une main nous disposons du savoir numérique, de l'autre d'une pratique de l'analogique, il ne nous reste plus qu'à devenir ambidextre.

Ce n'est sans doute pas un hasard si cet éditorial s'ouvre sur la suite de la transcription de nos Rencontres nationales des bibliothécaires musicaux à Paris. Celles-ci se déroulaient à la Bibliothèque nationale de France et à la Cité de la musique qui, par bien des aspects, illustrent ce que pourrait être une approche « bitextuelle » de la culture musicale.

**Musicalement vôtre,
Arsène Ott,
Président de l'ACIM**

Directeur de publication : Arsène Ott
Tél. 03 88 78 80 47 – Mél. ao.acim@gmail.com

Rédaction : Arsène Ott

**Transcription des interventions des Rencontres nationales :
Béatrice Pedot, de Marges et paperoles**

Crédits Photographiques : Arsène Ott

Légendes :

- Côté Atlantique ;

- Eléments de scénographie issus de l'exposition "Cadavres exquis" réalisée
par les Rhubarbus – Médiathèque de Strasbourg Centre ville.



Musique numérique :
oubli dans l'âme ou supplément d'âme ?

Sommaire :

Editorial :

**« Musique numérique : oubli dans l'âme ou supplément d'âme ? » /
Arsène Ott
(p. 1)**

Rencontres nationales des bibliothécaires musicaux, Paris 2009 -

« Le temps des cigales » :

**3e partie : transcription des interventions du début de l'après-midi du
lundi 23 mars 2009 à la BnF
(p. 6)**

Présidence de séance : *Elizabeth Giuliani*, adjointe à la directrice du département de l'Audiovisuel, BnF

**3.1 LE DÉPÔT LÉGAL DE L'INTERNET ET DES SITES MUSICAUX / *Gildas Illien*, chef
du service du dépôt légal numérique, département du dépôt légal, BnF, et *Alain Carou*,
chef du service images, département de l'audiovisuel, BnF (P. 6)**

**3.2 ARCHIVER LE NUMÉRIQUE : L'EXEMPLE DU MAGASIN NUMÉRIQUE DE LA BNF /
Thomas Ledoux et *Jérôme Dupont*, département des systèmes d'information, BnF
(P. 16)**

**3.3 ADAPTER LES CATALOGUES À L'ENVIRONNEMENT NUMÉRIQUE : RDA
(RESOURCE DESCRIPTION AND ACCESS) / *Françoise Leresche*, responsable du pôle
modélisation fonctionnelle, département de l'information bibliographique et numérique,
BnF (P. 25)**

**3.4 LE SON DANS GALLICA / *Pascal Cordereix*, chef du service des documents
sonores, département de l'audiovisuel, BnF (P. 37)**



Musique numérique : oubli dans l'âme ou supplément d'âme ?



3.1 LE DÉPÔT LÉGAL DE L'INTERNET ET DES SITES MUSICAUX

par **GILDAS ILLIEN**, chef du service du dépôt légal numérique, département du dépôt légal, BnF, et **ALAIN CAROU**, chef du service images, département de l'audiovisuel, BnF

- *Gildas Illien*, chef du service du dépôt légal numérique, département du dépôt légal, BnF

Nous allons faire une intervention à 2 voix avec Alain Carou. Pour ma part, je m'efforcerai de vous présenter le cadre général dans lequel nous mettons en œuvre le dépôt légal de l'Internet depuis bientôt 9 ans. J'insisterai sur la continuité que cette mission représente par rapport au travail de la Bibliothèque nationale de France (BnF) et j'essaierai de vous en dire plus – sans trop entrer dans la technique – sur le fonctionnement. Je passerai ensuite la parole à Alain Carou pour un zoom sur les collections musicales que nous avons pu collecter dans le cadre de cette activité nouvelle.

Archiver le web échappe un peu à la mesure de ce que nous pouvions imaginer. On change forcément de modèle compte tenu de sa volumétrie et, en même temps, nous restons dans une forte continuité des missions et des principes de l'enrichissement des collections patrimoniales de la BnF.

Le dépôt légal s'est adapté aux technologies à mesure qu'elles modifiaient les modalités de circulation des publications dans le domaine culturel et, plus largement, dans la société française. Il y a un effet presque mécanique d'adaptation de la loi et des pratiques patrimoniales de la bibliothèque aux innovations technologiques, mais il s'agit aussi, à chaque fois, de la prise de conscience que chacune de ces technologies modifie profondément les principes sociaux de circulation des publications. Nous avons l'exemple particulièrement pertinent, aujourd'hui, de la création d'un dépôt légal des partitions musicales en 1793 : c'est lorsqu'on se rend compte, à l'issue de la Révolution française, qu'elles ont aussi une portée politique particulière que le dépôt légal commence à s'effectuer de manière plus systématique. Au XIX^e siècle, la presse, avec l'alphabétisation et la création de l'école républicaine, prend une importance nouvelle dans la propagation des idées et c'est à cette époque que la bibliothèque renforce son effort pour la collecter systématiquement.

Nous pourrions faire l'analogie et dire que les années 2000, avec Internet, sont à la convergence de deux phénomènes : une innovation technologique qui appelle à une action patrimoniale et, en même

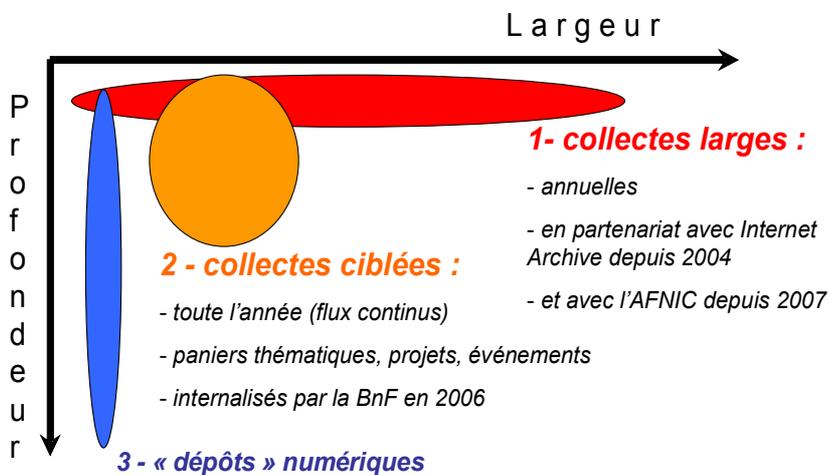


Musique numérique : oubli dans l'âme ou supplément d'âme ?

temps, un important changement de société, changement dont on ne peut pas encore dire exactement aujourd'hui où il nous conduit. Le web est à la fois un média de publication et un média de communication et nous devons essayer de le saisir dans toutes ses approches.

En 2006, le titre IV de la loi DADVSI, passé presque complètement inaperçu, officialisait un dépôt légal de l'Internet, dépôt légal que nous avons à cœur de mettre en application.

Comment avons-nous abordé la masse de l'Internet compte tenu du fait que, même si nous sommes une grande institution, nos moyens restent limités ? Nous avons choisi une approche qui combine une démarche de dépôt légal de très grande échelle, pour viser à garder la représentativité des usages du web à un moment donné, avec une approche plus documentaire dans le cadre de laquelle nous sélectionnons certains fonds, avec comme souci la continuité des collections patrimoniales mais aussi l'idée que la bibliothèque doit avoir un rôle prospectif pour sélectionner ce qui, sur la toile, mérite d'être conservé car significatif de l'évolution d'un média et de la manière dont il impacte les pratiques culturelles.



Sur ce schéma, les collectes larges (collectes automatiques du web) qui courent sur de très grandes quantités de données (plusieurs millions de fichiers) sont réalisées, depuis 2004, une fois par an (généralement à l'automne), en partenariat avec la fondation Internet Archive, basée en Californie. Depuis 2007, nous avons également un accord avec l'AFNIC² qui nous fournit tous les noms de domaine en .fr qui servent de point de départ à nos collectes par robot.

En complément de ces collectes exploratoires et à grande échelle, nous avons organisé dans la bibliothèque et avec des partenaires extérieurs, un réseau de meilleurs experts, qui réalisent la collecte

² Association Française pour le Nommage Internet en Coopération



Musique numérique : oubli dans l'âme ou supplément d'âme ?

ciblée de sites qui représentent un intérêt patrimonial particulier pour les collections de la BnF. C'est un réseau d'environ 80 bibliothécaires qui, dans tous les domaines, constituent des « paniers » de sites – dont un panier pour les collections audiovisuelles, de plus en plus conséquent en volume. Nous suivons plus spécifiquement certains événements (l'élection présidentielle de 2007 est un bon exemple d'événement générant, sur tous supports, des contenus éphémères mais extrêmement importants pour conserver la trace des événements nationaux). La bibliothèque combine ces deux approches et travaille aussi sur des formes plus expérimentales de dépôts pour des contenus qui échappent aux technologies existantes. C'est cette combinaison qui constitue notre dépôt légal, dans la continuité des collections imprimées ou sur d'autres supports.

Les notions de largeur et de profondeur sont importantes à comprendre. Plus vous collectez en profondeur, plus vous récupérez de publications complètes, mais plus cela coûte cher et plus c'est difficile. Plus vous collectez en largeur, plus étendue mais aussi plus superficielle est la collection. Ces collectes sont donc complémentaires : la collecte de surface sur beaucoup de sites, la collecte ciblée sur un nombre plus restreint de sites mais avec davantage de contenus.

Les robots de collecte

Un robot de collecte est un logiciel qui parcourt le web comme un internaute automatique. À partir d'une liste d'adresses URL, appelées graines (*seeds*), le robot copie et emmagasine les fichiers rencontrés, avec toutes les informations qui permettent de les relier aux autres. Si bien que ce qui est collecté constitue un « tissu » de liens : un ensemble navigable dans l'archive comme sur le web vivant.

Les technologies que nous utilisons, en particulier le robot Heritrix, logiciel libre développé par la communauté internationale qui s'intéresse à ces questions, sont des technologies déjà dépassées par le web. En conséquence, le robot rencontre régulièrement des obstacles et nos collections sont pleines de trous.

Les serveurs de collecte

Ce sont des serveurs comme ceux que vous avez dans vos bureaux mais nous en avons un certain nombre à partir desquels on lance des instances du robot en question.



Musique numérique : oubli dans l'âme ou supplément d'âme ?

La console de contrôle du robot Heritrix



C'est l'interface de suivi du robot. Chaque ligne représente un fichier qui est en train d'entrer dans la collection. On peut collecter un très grand nombre de documents en un temps record.

Tous ces fichiers sont stockés dans un format dédié aux archives du web : le format (W)ARC. C'est un format conteneur conçu pour emmagasiner très rapidement un grand nombre de fichiers.

Les petaboxes

Ces machines de stockage courant portent ce nom curieux car, après le téraoctet, le pétaoctet est désormais l'unité de mesure vers laquelle nous tendons pour les collections les plus vastes. Elles servent notamment à la consultation des archives, restreinte, pour des raisons juridiques, aux salles de lecture de la BnF. Ces machines font à peu près 2 mètres de haut et ont la taille d'un réfrigérateur américain.

Le système SPAR³

Il est utilisé pour l'archivage pérenne et la préservation. Nous avons une copie courante et une copie pour la préservation. Les données sont dupliquées dans ce magasin numérique de la bibliothèque.

L'accès aux informations

C'est aujourd'hui une de nos préoccupations essentielles. Nous indexons la totalité des archives pour permettre une navigation et une recherche par URL.

Nous expérimentons également l'indexation plein texte, qui n'est pas forcément la plus pertinente pour les documents musicaux. 20 % environ des collections peuvent actuellement être recherchées par mot : ce n'est pas le degré de performance d'un moteur de recherche comme Google mais nous travaillons de manière collaborative au niveau international sur des logiciels libres afin d'améliorer ces résultats.

³ Système de Préservation et d'Archivage Réparti



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Nous travaillons également sur d'autres modes de valorisation de ces fonds. Les volumétries en jeu sont telles, qu'il s'agit véritablement de trouver de nouveaux moyens pour sortir les aiguilles de la botte de foin.

L'interface de consultation proposée au public



La BnF conserve actuellement 12 milliards de fichiers : un fichier peut aussi bien être une petite vignette qu'un fichier son ou un gros PDF représentant l'équivalent d'une monographie.

L'ensemble de ces fichiers représente 136 téraoctets de données dont :

- des collections rétrospectives acquises auprès d'Internet Archive depuis 1996 ;
- des instantanés annuels du domaine français depuis 2004 ;
- des collectes ciblées réalisées à partir des sélections des bibliothécaires et de partenaires extérieurs, par exemple :
 - des collectes électorales (« web-campagnes ») : nous allons pour la première fois travailler sur la collecte des prochaines élections européennes en collaboration avec 7 bibliothèques nationales d'autres pays européens ;
 - des collectes des sites littéraires et d'écriture personnelle (vies numériques, journaux en ligne, blogs) ;
 - des collectes du web social et militant, par exemple la couverture des grèves, des mouvements sociaux ou des événements récents en outre-mer ;
 - des collectes de Dailymotion.

Ce sont des coups de projecteur sur certaines de nos collections.



Musique numérique : oubli dans l'âme ou supplément d'âme ?

■ *Alain Carou*, chef du service images, département de l'audiovisuel, BnF

Chaque service en charge de collections constitue, vous l'avez compris, des « paniers » de sites à collecter de manière approfondie. C'est le cas en musique, comme en histoire, en économie, en presse, en cinéma, etc.

La collecte ciblée dans le domaine de la musique

Elle concerne à la fois des contenus musicaux et des contenus relatifs à la musique et à la vie musicale :

- actualité musicale : sites d'informations, sites de festivals... ;
- documentation de l'édition musicale traditionnelle : sites des éditeurs phonographiques, des labels... ;
- édition musicale en ligne : plateformes de distribution de musique en ligne gratuite ou payante, « net labels » ; nous visons les contenus comme les fichiers sonores mais aussi la page dans laquelle s'inscrit ce fichier musical ;
- culture musicale au sens large : sites de critiques de disques, blogs et pages personnelles d'amateurs de musique et tout ce qui concerne les modalités de la culture musicale sur le web (discours sur le partage de musique, etc.).

Critères de la collecte ciblée

Cette sélection s'opère non pas selon une logique de la qualité informative ou esthétique des contenus mais selon une logique de représentativité propre au dépôt légal.

Elle consiste à repérer non seulement des sites de références qui font autorité mais aussi le tout venant de la production, y compris l'auto-publication qui, cependant, pour des raisons de masse évidentes, doit faire l'objet d'échantillonnages. Il n'est pas possible d'archiver toutes les pages MySpace, même en se restreignant aux pages en français !

Une partie de ce travail de sélection se trouve intégrée non seulement dans les tâches régulières des chargés de collections (en charge d'acquisitions, guidés par une logique de cohérence documentaire), mais elle entre aussi de plus en plus dans le travail des personnes qui sont en charge du dépôt légal, qui prospectent les publications et qui sont les mieux à même de repérer les continuités entre le dépôt légal traditionnel (sur supports) et le dépôt légal du web (par exemple, le passage de l'édition phonographique à la distribution en ligne, la publication uniquement en ligne des catalogues d'éditeurs, l'émergence du modèle des labels discographiques sur le web, avec les « net labels », etc.).



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Voici quelques exemples de sites qui font d'ores et déjà l'objet de collectes et pour lesquels nous disposons d'archives qui remontent sur plusieurs années :

- Les sites de culture musicale



Archive 2005 de La Blogothèque, site de critiques et de signalement de nouveautés dans le domaine musical.

- Les sites d'éditeurs phonographiques

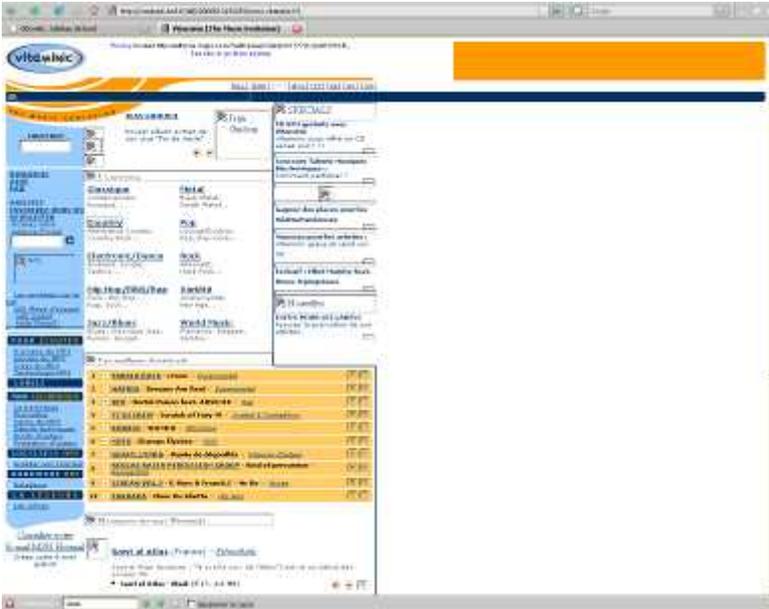


Archive 2006 de Mondomix.

Musique numérique : oubli dans l'âme ou supplément d'âme ?



- Les sites de musique en ligne



Archive 2000 de Vitaminic, premier site important de vente en ligne de musique. Ce site a aujourd'hui complètement disparu.

- Les sites de partage



Archive 2005 de Musique-libre.org, l'ancêtre de Dogmazic, site de partage de diffusion de musique libre et gratuite en Creative Commons.



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- Les blogs



Archive 2006 de Podaufeu, blog musical et site militant pour la musique libre. Il existe toujours aujourd'hui mais a changé d'adresse et, au passage, a perdu ses archives. Nous ne pouvons plus aujourd'hui remonter que sur quelques mois.



Archive 2001 de Central jazz, le blog d'un amateur de jazz.



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique : oubli dans l'âme ou supplément d'âme ?

Ce blog, très riche, permettrait d'étudier les goûts d'un amateur de jazz au fil du temps. Central jazz a disparu en 2006, apparemment sans trouver d'acheteur puisque, lorsqu'on tape centraljazz.com aujourd'hui, on est redirigé vers un site de photographies, géré par la même personne qui a changé de passion.

Limites et perspectives

Des limites spécifiques touchent les fichiers audio et vidéo en raison de la diversité des protocoles techniques et des empaquetages de données. Cette diversité rend la collecte plus compliquée que pour les contenus classiques des pages web. Néanmoins, des développements réguliers du robot nous permettent de résoudre un certain nombre de problèmes.

La piste des dépôts, c'est-à-dire la réclamation auprès des éditeurs des contenus de fichiers que nous n'arrivons pas à recueillir avec le robot, devra être explorée davantage dans les années à venir mais elle requerra une réglementation spécifique. Elle nous permettra une extension aux contenus payants qui restent aujourd'hui complètement en dehors du périmètre de la collecte.

Des collections accessibles

Les collections ainsi constituées sont accessibles dans les salles de recherche de la BnF : il ne s'agit donc pas – vous l'avez compris – de l'ensemble du web, mais essentiellement du domaine français, construit par échantillonnage ; ce sont là des collections fort riches qui représentent, par les outils de navigation disponibles, une source d'information d'appoint très précieuse par rapport aux sources traditionnelles. C'est sans doute aussi un outil pour des études construites et centrées autour du web.

Des méthodes et des outils restent à inventer pour avancer dans la recherche sur ce type de collections et nous sommes à l'écoute des demandes des chercheurs pour que leur accès soit plus efficace à l'avenir.

Questions de la salle

■ *Arsène Ott*, président de l'ACIM

Pourriez-vous préciser la collaboration avec Internet Archive ?

■ *Gildas Illien*

Internet Archive est une fondation à but non lucratif qui dès 1996 a commencé à archiver le web mondial. En 2004, la BnF a signé un accord de recherche et de développement avec Internet Archive qui consistait à ce qu'Internet Archive réalise pour la BnF les collectes de grande échelle, en échange



Musique numérique : oubli dans l'âme ou supplément d'âme ?

de quoi la BnF apportait son expérience de bibliothèque nationale et finançait une partie des outils utilisés dans le cadre d'un consortium international pour la préservation de l'Internet.

Cinq ans plus tard, nous considérons que la BnF a atteint une certaine maturité et qu'elle est en mesure d'internaliser ces collectes larges, mais elle continuera à collaborer avec Internet Archive pour le développement des logiciels.



3.2 ARCHIVER LE NUMÉRIQUE : L'EXEMPLE DU MAGASIN NUMÉRIQUE DE LA BNF

par *THOMAS LEDOUX* et *JÉRÔME DUPONT*, département
des systèmes d'information, BnF

Nous présenterons ici l'ensemble des collections numériques que peut traiter et héberger la BnF, le traitement de l'archivage numérique pour des collections patrimoniales, la solution que nous sommes en train de mettre en œuvre et enfin un focus sur la vie d'un document numérique.

LES COLLECTIONS NUMÉRIQUES DE LA BNF

Le contexte

Il est très varié et pose des problématiques différentes :

- La production documentaire uniquement sous forme numérique : c'est par exemple le dépôt légal du web ou le dépôt légal « négocié » d'organes de presse. Ce sont des fichiers numériques « natifs » que nous devons héberger et préserver.
- La dématérialisation des supports analogiques : pour certains d'entre eux, le support original a disparu et il est important de réaliser des opérations de numérisation de masse pour les préserver.
- La numérisation suite à l'obsolescence voire la disparition des équipements de restitution : c'est le cas notamment de la numérisation de sauvegarde des collections audiovisuelles.
- La disparition progressive des moyens de production de microformes : elle nécessite le remplacement des moyens de capture analogique.



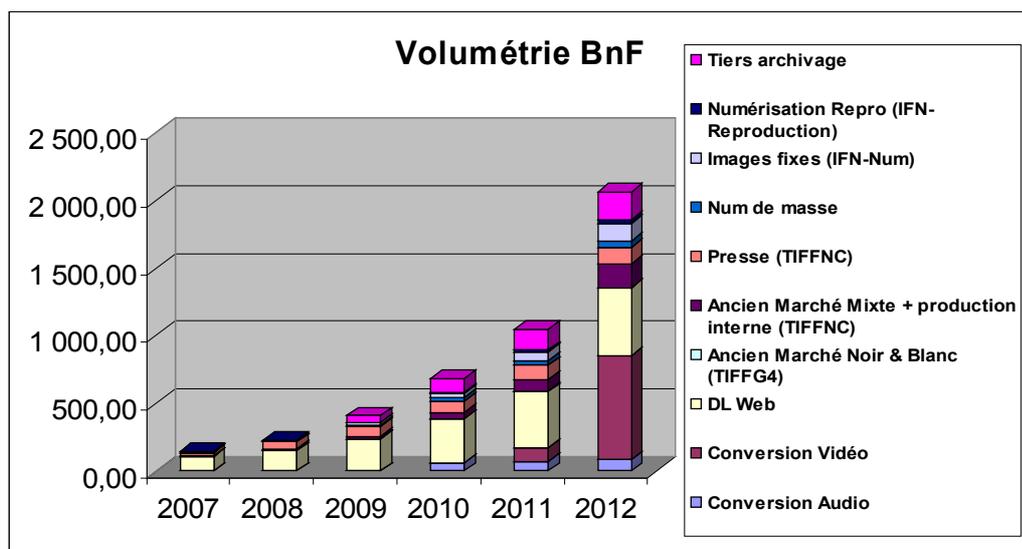
Musique numérique : oubli dans l'âme ou supplément d'âme ?

État des lieux des collections

- Diversité de supports et des formats d'enregistrement : la variété d'informations alimente un fonds important avec beaucoup de supports physiques différents à manipuler. Pour s'abstraire de cette contrainte, ces supports sont dématérialisés ;
- Diversité de formats de fichiers/de structures (TIFF, JPEG, MPEG, HTML, XML, etc.) ;
- Diversité des environnements matériels et logiciels qui dépendent du type de documents à diffuser ;
- Diversité des durées de vie très variables les unes des autres.

Nous cherchons à mutualiser toutes ces contraintes.

Volumétrie des collections à traiter



Aujourd'hui, le cumul de l'ensemble des collections est encore en dessous du pétaoctet, mais nous allons rapidement atteindre ce seuil et le doubler dans les années à venir. Il nous faut donc maîtriser la croissance exponentielle et mettre en œuvre les moyens nécessaires pour collecter, héberger et diffuser toute ces informations. C'est le véritable défi de l'archivage numérique.

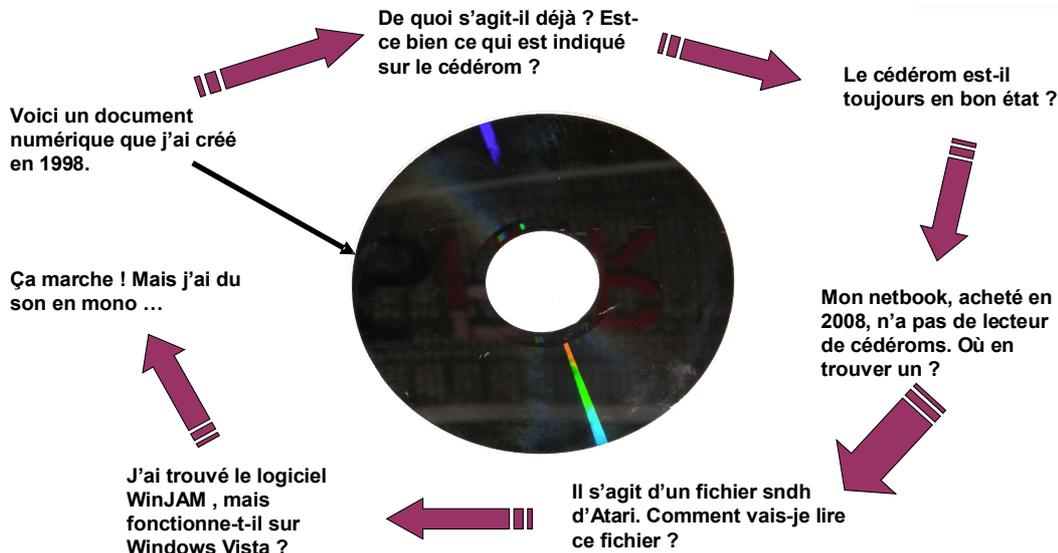
L'ARCHIVAGE NUMÉRIQUE

Préservation et numérique

Prenons l'exemple d'un CD, créé en 1998 et pour lequel nous devons pouvoir relire ce qui y est inscrit.



Musique numérique : oubli dans l'âme ou supplément d'âme ?



Sur cet exemple, il s'agit d'un jeu, terrain en constante évolution, et pour lequel nous avons des difficultés pour préserver la connaissance qui permet de lire et comprendre le contenu, mais ce problème se rencontre pour tous les documents numériques.

Il faut pouvoir préserver les données en s'assurant de ne pas avoir abîmé le support sur lequel on peut lire les informations mais, *in fine*, nous devons nous assurer que, dans les années à venir, une communauté définie d'utilisateurs saura lire et comprendre les informations qui figurent dans un document numérique donné : il s'agit donc de préserver toutes les informations qui permettent de comprendre ce document.

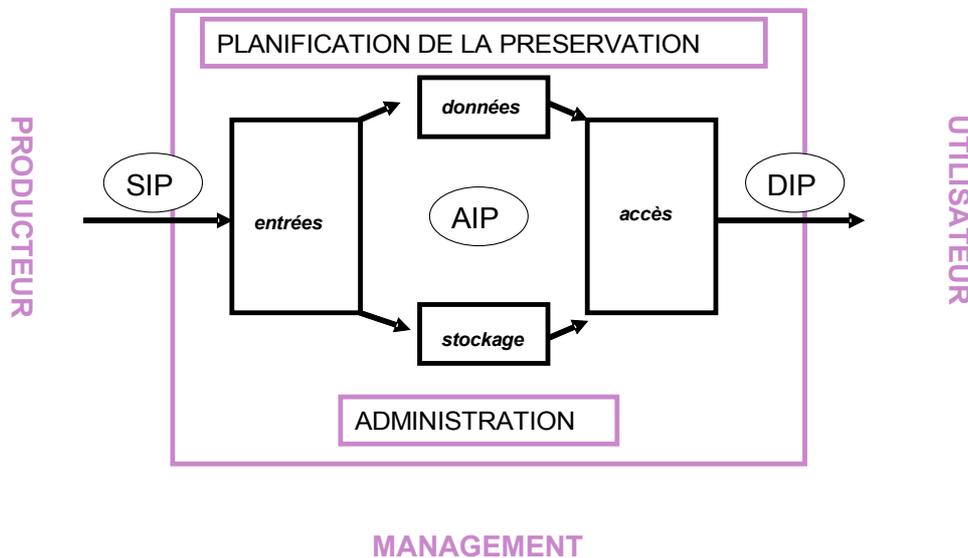
Un modèle normatif OAIS (Open Archival Information System) – ISO 14721

Cette norme définit la notion de paquet d'informations qui contient l'information elle-même (le document, le jeu, l'enregistrement sonore) et toutes les informations qui permettront de décoder, de lire et de comprendre cette information.

Elle définit le cycle de vie du paquet.



Musique numérique : oubli dans l'âme ou supplément d'âme ?



SIP = *Submission Information Package* (paquet de versement). Au moment où le document entre, nous vérifions si nous arrivons à le lire et le comprendre et nous y adjoignons toutes les données qui permettront de le comprendre dans le futur.

Ces paquets se transforment en paquets d'archives (AIP) qui seront stockés sur une longue période et pour lesquels il faut s'assurer qu'ils pourront être lus et compris au fil des années.

Enfin, le jour où l'utilisateur veut lire ce paquet d'informations, il est disséminé (DIP) et l'information est présentée à l'utilisateur sous la forme qu'il désire.

LA SOLUTION MISE EN ŒUVRE A LA BNF

Les filières et leurs chaînes

Nous avons vu qu'il y avait, d'une part, une énorme quantité d'informations à gérer et que, d'autre part, une norme a été élaborée. Il faut donc mettre en relation ces deux notions pour obtenir une solution viable et utilisable.

La première chose que nous avons essayé de définir est la manière d'ordonner ces paquets de fichiers numériques à gérer. Nous avons défini des filières pour hiérarchiser les différentes opérations à réaliser sur ces fichiers et, en particulier, selon la latitude que nous pouvons avoir à leur égard :

- numérisation de conservation ;
- audiovisuel ;
- numérisation de consultation ;
- dépôt légal automatique ;



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- dépôt légal négocié ;
- production administrative/technique ;
- dons et acquisitions ;
- tiers archivage.

Par exemple, pour les fichiers provenant du dépôt légal, nous n'avons aucune latitude par rapport aux fichiers entrants : nous collectons à partir du web et nous ne pouvons pas modifier le fichier lui-même pour des besoins de conservation.

A contrario, lorsque nous faisons de la numérisation de conservation, nous décidons du procédé de numérisation et nous pouvons choisir des formats que nous maîtrisons bien.

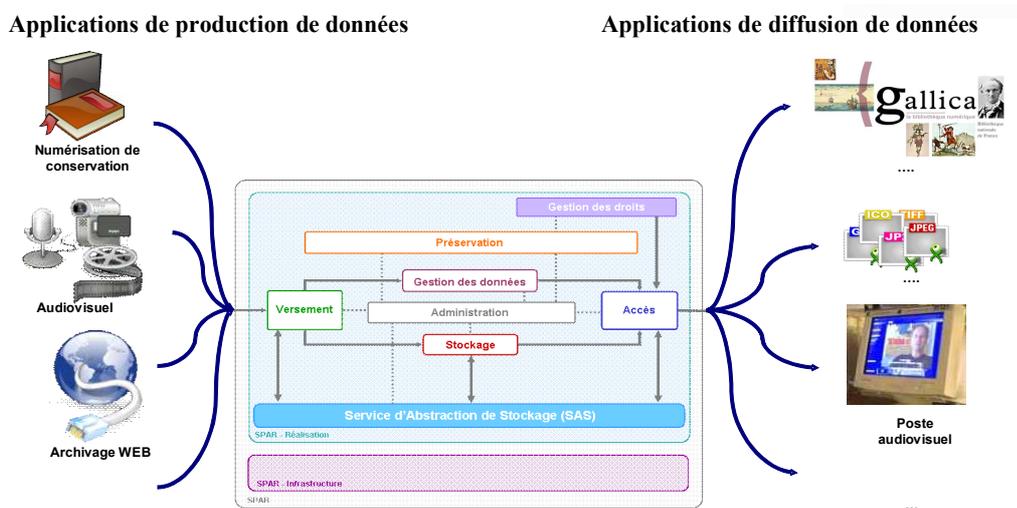
Selon la manière dont nous pouvons agir sur les fichiers, nous mettons en place des stratégies de préservation différentes soit par migration de format, soit, au contraire, par émulation.

Il est important d'avoir ces filières comme trame de décision par rapport à la préservation.

Le système SPAR (Système de préservation et d'archivage réparti)

Ce système est au cœur de l'ensemble des opérations de manipulation de l'information numérique de la BnF.

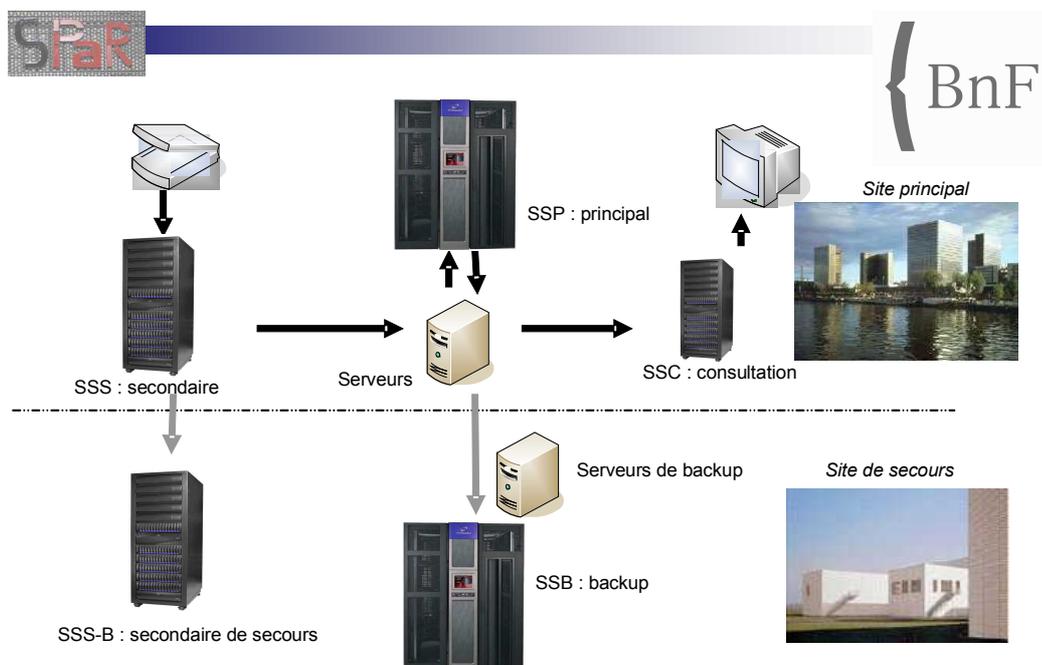
Il repose quasi exclusivement sur la norme OAIS qui permet d'avoir un canevas et de ne rien oublier. Nous retrouvons dans ce système l'ensemble des entités OAIS, auxquelles nous avons été obligés d'ajouter un module particulier, celui de la gestion des droits, ainsi qu'un module plus technique d'abstraction de stockage qui permet de continuer à utiliser le même système quelle que soit l'infrastructure mise en dessous.





Musique numérique : oubli dans l'âme ou supplément d'âme ?

L'infrastructure



Le système est hébergé sur deux sites : le site principal à Tolbiac et un site de secours qui nous permet d'avoir un plan de reprise d'activités en cas d'incident grave sur le site principal. Nous retrouvons les trois moments de l'information :

- le stockage secondaire permet de récupérer les informations en versement ;
- une fois ces informations traitées, elles sont hébergées dans de grandes bandothèques pour être conservées sur le long terme ;
- en cas de besoin, au moment de la consultation, on génère des exemplaires de diffusion, mis sur disques pour accroître la rapidité.

LA VIE D'UN DOCUMENT NUMÉRIQUE

Les métadonnées

Quelles sont les métadonnées ajoutées au document pour permettre de le lire dans le futur ?

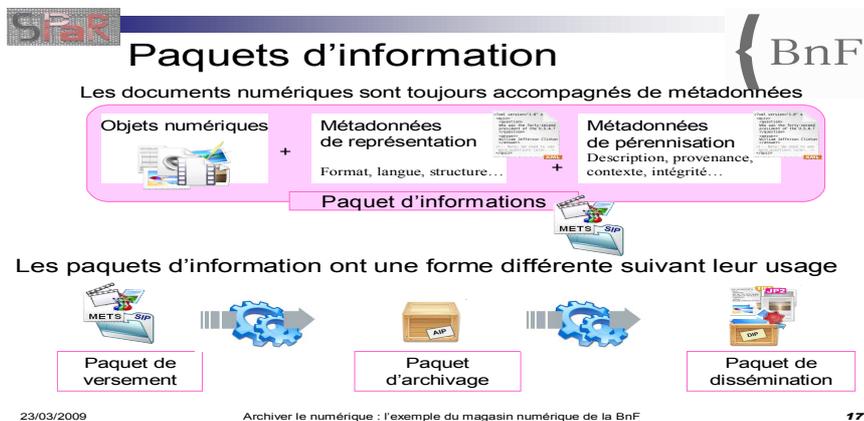
- Métadonnées descriptives :
 - Le titre : L'autre monde, ou les estats & empires de la lune
 - Les auteurs : Cyrano de Bergerac, Savinien de, Dufaut, François,...



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- La date de publication : DL 2005
- L'éditeur : Alpha
- ...
- Métadonnées techniques :
 - En-têtes des fichiers audio *Wav 16bits 44kHz*
- Métadonnées de structure :
 - Fichier principal
 - Synopsis
- Métadonnées de source :
 - Ce document a été numérisé à partir du cédérom SDC12_197362
- Métadonnées d'historique :
 - Le document a été numérisé le 18/09/2008 avec un PHILIPS DVD8631
 - Le document a été stocké sur bande le 18/09/2008

Les paquets d'information



Les métadonnées de représentation permettent de lire le contenu d'un document.

Les métadonnées de pérennisation permettent de comprendre l'historique des actions accomplies sur le document depuis qu'il est

entré dans l'archive.

Le paquet de versement est le paquet tel qu'il arrive chez nous : il contient essentiellement l'objet donné et une description sommaire.

Dans le paquet d'archivage, on ajoute les métadonnées techniques qui permettent de continuer à le lire.

Le paquet de dissémination offre une vue de ce paquet, éventuellement transformé : par exemple, pour une vue imprimée, on fournira une résolution moins importante que celle retenue pour l'archive.



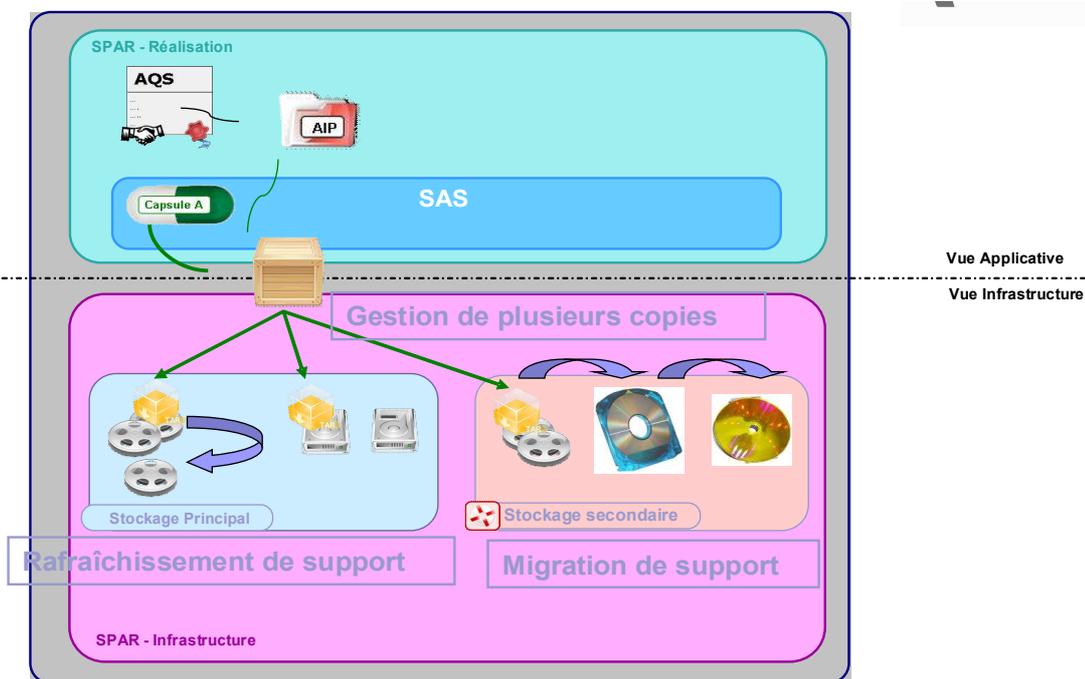
Musique numérique : oubli dans l'âme ou supplément d'âme ?

L'archivage sur l'infrastructure

Lorsqu'un paquet entre dans le système, il faut s'assurer de sa survie et, pour cela, nous réalisons ainsi une copie en plusieurs exemplaires à son entrée. Si une copie est endommagée, nous avons toujours la capacité de lire la seconde. Un contrôle régulier est effectué pour éviter que les deux copies ne s'abîment en même temps.

Lorsque le stockage se fait sur bandes, nous pouvons également les contrôler en opérant un rafraîchissement de support qui consiste à le recopier sur une autre cassette de même technologie avant qu'il ne devienne illisible.

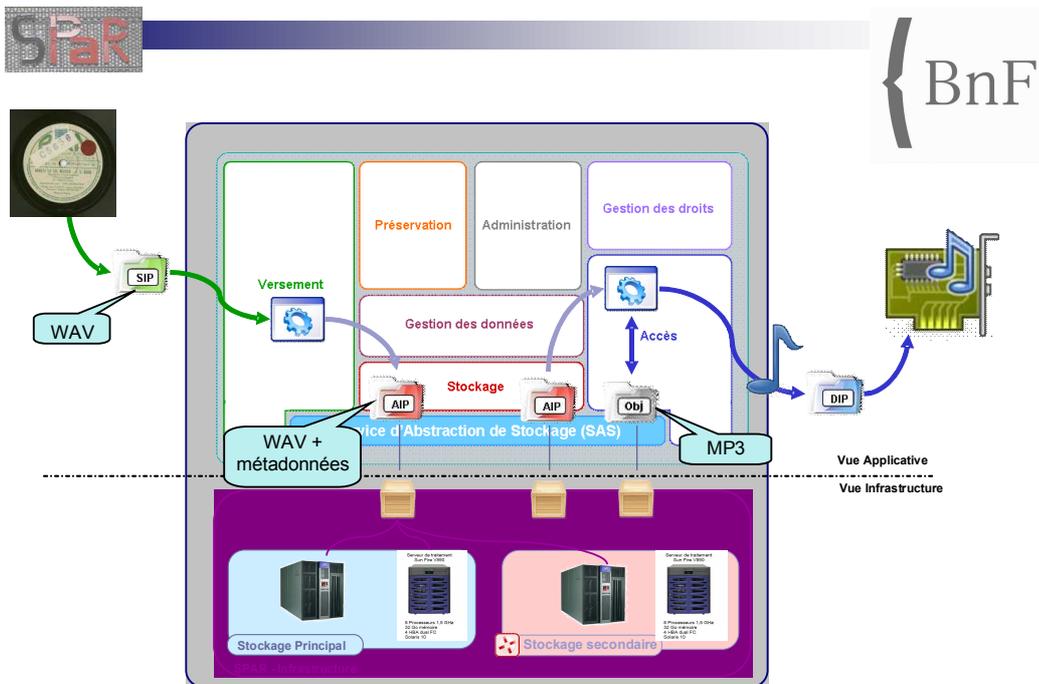
Les technologies de stockage informatique évoluant très vite, des migrations de support consistent à changer de technologie tous les 5 ou 10 ans pour avoir de plus grands volumes ou une plus grande fiabilité.





Musique numérique : oubli dans l'âme ou supplément d'âme ?

EN RESUMÉ



Le document arrive à la BnF, *via* le module versement, où il subit tous les contrôles nécessaires. On ajoute les métadonnées et on le transforme en paquet d'archive. Il est alors copié sur deux sites différents. Pour y accéder, on sort le paquet de l'archive, il est éventuellement transformé en fonction de la demande de l'utilisateur et transmis à l'utilisateur.

Questions de la salle

■ *Une personne du public*

Faites-vous en sorte que les documents soient lisibles sur tous les OS (systèmes d'exploitation) existants ?

■ *Thomas Ledoux*

Nous nous occupons essentiellement des problèmes de format des fichiers. C'est pourquoi nous distinguons les trois moments de l'archivage des paquets. Nous cherchons à mettre dans les AIP des formats aussi standards que possible. Il y a des filières pour lesquelles nous pouvons imposer un standard : lorsque nous numérisons, nous choisissons des standards aussi ouverts que possible. Mais



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique : oubli dans l'âme ou supplément d'âme ?

pour d'autres, nous n'avons pas de maîtrise des formats : c'est typiquement le cas pour le web généré par un producteur à qui nous ne pouvons imposer un langage. Nous adaptons alors plutôt des techniques d'émulation.

Les DIP, eux, seront disséminés pour une technologie à un instant donné. Aujourd'hui, par exemple, pour diffuser du son, nous utilisons du MP3. Mais le MP3 ne sera pas conservé, car la technologie va encore s'améliorer et changer, il sera donc nécessaire d'avoir d'autres moyens de diffusion.

Nous distinguons toujours les paquets d'archive, qui doivent être durables et au format le plus standard possible, des paquets de diffusion, beaucoup plus volatiles et non voués à être archivés en tant que tels.



3.3 ADAPTER LES CATALOGUES À L'ENVIRONNEMENT NUMÉRIQUE : RDA (RESOURCE DESCRIPTION AND ACCESS)

par *FRANCOISE LERESCHE*, responsable du pôle
modélisation fonctionnelle, département de l'information
bibliographique et numérique, BnF

Cette communication a pour but de vous présenter les évolutions en cours dans le domaine du catalogage, évolutions qui sont largement motivées par l'environnement numérique et le foisonnement des ressources numériques que nous connaissons actuellement.

LE CONTEXTE

Un environnement en mutation

- **Des documents plus complexes**

Les ressources électroniques ont un contenu plus riche que les documents traditionnels, avec une organisation en profondeur plus complexe. Les sites web sont typiques de cette organisation en profondeur.

Par ailleurs, les sites web et les bases de données posent de façon différente la question de la « sérialité », car ce sont des ressources qui gardent le même contour, mais s'enrichissent ou sont modifiées chaque jour : on a toujours affaire à **une** ressource, mais dont le contenu n'a pas de forme



Musique numérique : oubli dans l'âme ou supplément d'âme ?

fixe – ce qui est très différent des publications en série « traditionnelles » qui sont constituées par un ensemble de fascicules ou de volumes qui s'ajoutent les uns aux autres selon une certaine périodicité.

Enfin, les ressources électroniques posent le problème de l'articulation entre le support de publication et le contenu. Tout catalogueur amené à cataloguer des ressources électroniques – et particulièrement à partir des normes actuelles – se pose la question de ce qu'il doit décrire en priorité : la ressource électronique ou son contenu (qui peut être du texte, de la musique ou des images).

Depuis une dizaine d'années, on assiste également à la multiplication des formes de présentation, en particulier avec des publications parallèles en ligne et sur support physique, avec des contenus variables selon la forme de présentation qui est choisie (les sites de journaux en sont un exemple).

Du fait de toutes ces caractéristiques, il est de plus en plus important de pouvoir relier les ressources décrites.

• Multiplication des métadonnées

On assiste à un développement des métadonnées et de leurs usages. Les descriptions bibliographiques telles qu'elles sont produites par les bibliothèques intéressent de plus en plus des acteurs extérieurs aux bibliothèques, comme les circuits de l'édition et de la diffusion, mais aussi les gestionnaires de droits.

Parallèlement, avec la diffusion d'Internet, la plupart des communautés professionnelles se dotent de schémas de métadonnées spécifiques, tandis que des schémas de métadonnées plus génériques comme le Dublin Core sont créés en vue d'assurer l'interopérabilité sur le web.

• Évolution de la structuration de l'information

Elle est essentiellement liée à Internet, avec de nouveaux formats :

- le langage XML (*eXtensible Markup Language*) ;
- un nouveau langage émergent, le langage RDF (*Resource Description Framework*) qui est le langage du web sémantique.

• Nouvelle approche de l'interopérabilité

Toutes ces évolutions ont permis une nouvelle vision de l'interopérabilité : il est désormais possible de définir des applications spécifiques pour répondre à des besoins particuliers. Des liens peuvent être établis entre ces données hétérogènes, en utilisant des ontologies ou des modèles de référence comme le modèle CIDOC-CRM⁴ dans le domaine patrimonial. Le projet VIAF (*Virtual International Authority*

⁴ Comité International pour la DOcumentation du conseil international des musées – Conceptual Reference Model



Musique numérique : oubli dans l'âme ou supplément d'âme ?

File) qui établit des correspondances entre les notices d'autorité de pays différents est un exemple de cette nouvelle conception de l'interopérabilité.

Un nouveau cadre théorique

Face à ces mutations, les fondements théoriques, eux aussi, ont évolué. Le rapport final sur les FRBR (*Functional Requirements for Bibliographic Records* = Fonctionnalités requises des notices bibliographiques), paru en 1998, contient un modèle conceptuel de l'information bibliographique qui formalise le contenu de l'information gérée dans les catalogues et permet de l'organiser de manière fructueuse.

Prolongement des FRBR

Le modèle FRBR est à la base de toutes les réflexions actuelles sur les règles de catalogage. Il a par ailleurs connu des prolongements :

- la modélisation des notices d'autorité, avec le modèle FRAD (*Functional Requirements for Authority Data*) à paraître en 2009 ;
- la modélisation de la relation de sujet, avec le modèle FSAR (*Functional Requirements for Subject Authority Records*) en cours d'élaboration au sein de l'IFLA⁵.

En liaison avec la nouvelle conception de l'interopérabilité évoquée précédemment, l'IFLA a souhaité rapprocher le modèle FRBR d'autres modèles conceptuels, en particulier le modèle CIDOC-CRM, développé au départ par la communauté des musées mais dont le domaine d'application est plus large. Ce modèle est d'ailleurs devenu une norme ISO (ISO 21127), sous le titre *Une ontologie de référence pour l'échange d'informations du patrimoine culturel*⁶. Il s'agit d'un modèle orienté objet. Le rapprochement du modèle FRBR avec le modèle CIDOC-CRM a donc conduit à la définition d'une version orientée objet du modèle FRBR, FRBRoo. Cette dernière précise les entités du modèle FRBR, codifie les relations et les exprime sous forme de « triplets » RDF : elle facilite donc la conversion du modèle vers le formalisme de RDF et devrait permettre aux catalogues de bibliothèques d'évoluer vers les applications du web sémantique.

La révision des *Principes de Paris*

Jusqu'en 2009, les *Principes de Paris*, qui datent de 1961, constituaient encore la référence en matière de règles de catalogage, bien qu'ils soient devenus largement inadaptés au contexte actuel des catalogues : ils étaient conçus pour des catalogues sur fiches, ne contenant que des documents textes

⁵ International Federation of Library Associations and Institutions <http://www.ifla.org/>

⁶ ISO 21127 : 2006, version anglaise 5.0.1, mars 2009 : http://cidoc.ics.forth.gr/docs/cidoc_crm_version_5.0.1_Mar09.pdf



Musique numérique : oubli dans l'âme ou supplément d'âme ?

imprimés, et plus particulièrement pour les catalogues par auteurs et titres d'anonymes (ils ne traitent que les entrées auteurs/titres).

Un travail de révision de ces principes a été entrepris par l'IFLA pour les adapter au nouvel environnement des catalogues : désormais les catalogues sont informatisés et accessibles sur le web ; ils sont multimédia et la documentation électronique y tient une place de plus en plus importante. Par ailleurs, les exigences du public vis-à-vis des catalogues ont également changé.

L'objectif de ce travail était de définir de nouveaux principes de catalogage qui puissent servir de base à un code international de catalogage pour le XXI^e siècle. Le groupe qui a présidé à ces travaux au sein de l'IFLA avait d'ailleurs pour sigle IME-ICC (*IFLA Meeting of Experts for an International Cataloguing Code*).

Le travail a été échelonné sur 5 ans (2003-2007). Il s'est déroulé sous la forme d'un cycle d'ateliers régionaux qui, à chaque fois, se concentraient sur une région du globe avec une cohérence catalographique forte (Europe en 2003, Amérique latine en 2004, Moyen-Orient en 2005, Extrême-Orient en 2006 et Afrique sub-saharienne en 2007). Chaque atelier réunissait des experts de l'IFLA et des responsables des organismes nationaux chargés de la normalisation du catalogage.

Ce travail a abouti, en 2009, à la publication de nouveaux *Principes internationaux de catalogage*⁷. Les évolutions rapides du contexte technique des catalogues, comme les travaux en cours dans le domaine de la modélisation, ont conduit l'IFLA à définir également le processus d'actualisation de ces nouveaux principes, afin de pouvoir les faire évoluer au fur et à mesure des besoins.

Principales caractéristiques des Principes internationaux de catalogage

- Ils sont élaborés pour des catalogues en ligne ;
- Ils couvrent tous les types de documents ;
- Ils traitent de tous les aspects des catalogues : des notices bibliographiques, mais aussi des notices d'autorité ;
- Ils prennent en compte tous les points d'accès :
 - par sujets (vedettes-matière, classification) ;
 - à partir d'éléments de la description bibliographique (identifiant normalisé, date, etc.) ;
 - en utilisant tous les mécanismes de recherche disponible dans les interfaces des catalogues en ligne (recherche par mots, troncature) ;

⁷ Traduction française : www.ifla.org/VII/s13/icp/ICP-2009_fr.pdf



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- Ils définissent les fonctions d'un catalogue à partir des tâches des utilisateurs telles qu'elles sont listées dans le rapport final sur les FRBR (trouver, identifier, sélectionner, se procurer). Le confort de recherche de l'utilisateur du catalogue est affirmé comme le premier des principes.
- Ils s'appuient expressément sur les entités définies par les FRBR ;
- Ils listent les objectifs qui doivent guider l'élaboration d'un code de catalogage, à savoir les principales qualités attendues des règles de catalogage ;
- Ils se caractérisent par une évolution du **vocabulaire** qui correspond au contexte électronique des catalogues : ainsi, on ne parle plus de « vedettes » mais de « points d'accès ». Au fur et à mesure de l'élaboration de ces Principes, de 2003 à 2008, on parle de moins en moins de « notices », mais beaucoup plus de « données » ou d'« ensemble de données », pour viser une intégration de plus en plus forte des catalogues sur le web et, en particulier, sur le web sémantique.

UN NOUVEAU CODE DE CATALOGAGE : RDA (RESOURCE DESCRIPTION AND ACCESS)

Des AACR à RDA

Ce nouveau code de catalogage a été conçu pour l'environnement numérique.

Il s'inscrit dans une histoire relativement longue car il est l'héritier des règles de catalogage anglo-américaines (AACR, *Anglo-American Cataloguing Rules*).

C'est en 2003 que l'organisme chargé de la maintenance des AACR, le Joint Steering Committee for the revision of AACR (JSC), a pris la décision de réviser en profondeur les AACR pour :

- prendre en compte l'apport des FRBR ;
- adapter le code de catalogage au contexte actuel (électronique) des catalogues ;
- le rendre conforme aux *Principes internationaux de catalogage* en cours d'élaboration.

En 2005, il décide d'aller encore plus loin et d'élaborer un nouveau code de catalogage qui remplacera les AACR, RDA (*Resource Description and Access*). La publication de ce code est annoncée pour le 3^e trimestre 2009⁸.

⁸ Présentation sur le site du JSC for development of RDA : www.collectionscanada.gc.ca/jsc/rda.html



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Les objectifs de RDA

- **Fournir des règles :**
 - pour décrire tous les types de ressources, les ressources électroniques comme les ressources traditionnelles ;
 - fondées sur les *Principes internationaux de catalogage* ;
 - faisant référence aux modèles de l'information bibliographique (FRBR et FRAD) pour permettre de résoudre la problématique de la description du contenu **et** de sa présentation matérielle sur un support particulier ;
 - compatibles avec les normes reconnues au niveau international, dont l'ISBD, le rôle de celui-ci ayant été réaffirmé par les *Principes internationaux de catalogage*.

- **Produire des notices :**
 - capables de répondre aux besoins des utilisateurs ;
 - compatibles avec les notices présentes dans les catalogues actuels ;
 - indépendantes du format et du système utilisé pour stocker ou communiquer les données : l'accent est mis sur l'information requise pour décrire une ressource et non sur la présentation de cette information ;
 - adaptables dans les structures émergentes de bases de données pour préparer l'évolution technique des catalogues ;
 - utilisables dans l'environnement du web.

- **Être un code de catalogage :**
 - à vocation internationale ;
 - utilisable par tous les types de bibliothèques, de la plus grande à la plus petite, de la bibliothèque de lecture publique à la bibliothèque à vocation patrimoniale ;
 - ouvert à d'autres communautés que la communauté des bibliothèques (archives, documentation) ;
 - adapté à l'environnement actuel du travail des catalogueurs : un outil en ligne qui prendra la forme d'un site web.

Les caractéristiques de RDA

La référence aux modèles de l'information bibliographique est extrêmement étroite :

- modèle FRBR pour l'information bibliographique ;



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- modèle FRAD pour les données d'autorité (contenu des notices d'autorité pour les personnes, les œuvres, les collectivités).

Le code de catalogage va s'attacher à rendre compte des entités FRBR et FRAD avec leurs principales caractéristiques (leurs attributs) et les relations qu'elles peuvent avoir entre elles. Chaque élément de données est associé, d'une part, aux entités FRBR/FRAD et, d'autre part, aux tâches des utilisateurs du catalogue.

L'organisation générale est calquée sur le modèle FRBR avec 10 sections réparties en 2 grands groupes qui vont :

- décrire les **attributs** des entités :
 - manifestation / document ;
 - œuvre / expression ;
 - personne / collectivité / famille ;
 - concept / objet / événement / lieu.
- décrire les **relations** entre ces entités :
 - entre oeuvre / expression / manifestation / document ;
 - entre une ressource et les personnes, collectivités, familles ;
 - entre une œuvre et les concepts, objets, évènements, lieux ;
 - entre des œuvres, expressions, manifestations, documents ;
 - entre des personnes, collectivités, familles ;
 - entre des concepts, objets, événements, lieux.

La référence au modèle FRBR permet de décrire de manière satisfaisante les différents aspects d'une ressource, en particulier l'articulation entre le contenu et le contenant :

- type et forme du **contenu** :
 - forme → attribut de l'**œuvre** (exemples : poème, film, symphonie...);
 - type → attribut de l'**expression** (exemples : parole, son, interprétation musicale, musique notée...).
- type et forme de la **présentation matérielle** :
 - attributs de la **manifestation** avec deux niveaux de précision :
 - générique (exemples : pas de médiation, audio, vidéo, électronique...);
 - spécifique (exemples : feuille, volume... ; disque audio, cassette audio, bande... ; ressource en ligne, disque, disquette...).



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Par ailleurs, les règles définies par RDA permettent de s'inscrire dans un contexte numérique, en offrant la possibilité de récupérer automatiquement les données et de s'affranchir des règles de présentation (ponctuation...). Elles s'affranchissent des limites imposées par le format d'une fiche et permettent la multiplication des points d'accès, si c'est jugé utile pour les utilisateurs du catalogue.

RDA se présente comme une norme de contenu qui s'attache à définir l'information requise pour décrire une ressource :

- **il traite du contenu** de l'information bibliographique, c'est à dire des éléments d'information et de leurs relations avec les entités FRBR et avec les tâches des utilisateurs ;
- **il s'affranchit** des modes de **présentation** qui sont seulement traités sous forme d'annexes. C'est le cas pour l'ISBD qui est assimilé à un format d'affichage et de la correspondance en format (MARC 21 et Dublin Core).

RDA vise à une évolution de la structure des catalogues pour permettre leur intégration dans le web sémantique et cherche à rapprocher la structuration de l'information des modèles orientés objet et des structures relationnelles. Cette transition sera progressive avec deux étapes prévues :

- un scénario à court terme avec le maintien de la structure actuelle des catalogues basée sur 3 fichiers (notices bibliographiques, données d'autorité et notices d'exemplaires) et une évolution des formats MARC pour permettre de rendre compte de toutes les entités FRBR ;
- un scénario à moyen terme qui sera une structure reflétant le modèle FRBR avec des notices pour les entités **œuvre** et **expression**.

VERS UN CODE INTERNATIONAL DE CATALOGAGE

Les atouts de RDA

RDA a tous les atouts pour devenir le futur code de catalogage que les *Principes internationaux* doivent permettre d'élaborer. Il est :

- complet et couvre les données bibliographiques et d'autorité, les points d'accès, les relations entre les données ;
- conforme aux *Principes internationaux de catalogage* ;
- d'utilisation « supranationale » de fait puisque toute la communauté qui utilisait les AACR va passer à RDA ;
- élaboré en consultant les autres traditions catalographiques européennes (Allemagne, France, Espagne, Suède, etc.).



Musique numérique : oubli dans l'âme ou supplément d'âme ?

RDA et l'IFLA

La Section de Catalogage de l'IFLA soutient RDA : du fait de la forte présence nord-américaine au sein de la section, mais aussi par l'organisation de séminaires sur RDA, par la publication d'une annonce sur la mise en œuvre de RDA par les quatre bibliothèques nationales concernées au premier chef (British Library, Library of Congress, Bibliothèque et Archives nationales du Canada, National Library of Australia).

On peut affirmer que RDA deviendra le futur code international de catalogage. La seule incertitude porte sur la date d'entrée en vigueur. L'IFLA se gardera un temps d'observation pour examiner la réalité de l'application du nouveau code dans les bibliothèques nationales concernées, mais ce délai ne devrait pas être trop long.

RDA en France aussi ?

Que faire avec les normes AFNOR lorsque RDA sera devenu le code international de catalogage ? Faut-il remplacer les normes AFNOR par RDA ?

La question est ouverte car :

- nous n'avons rien d'aussi solide à opposer : peu de forces se consacrent en France à la normalisation du catalogage ;
- l'IFLA soutient RDA ;
- RDA est un code qui a été élaboré dans le cadre d'une réflexion bibliographique internationale novatrice et stimulante et qui présente des orientations extrêmement intéressantes, dans la perspective de la transition des catalogues vers le web sémantique.

Une réflexion nationale sera nécessaire pour étudier l'opportunité d'adopter RDA en France et les modalités de cette implémentation. Elle ne pourra pas débuter avant que la version stabilisée du nouveau code soit publiée, ce qui est annoncé pour le 3^e trimestre 2009. En conséquence, un groupe national sera constitué en 2010 pour réfléchir à ces questions.

Les enjeux de demain

Aujourd'hui, les catalogues sont sur le web mais leurs données ne sont pas exploitables par les moteurs de recherche et ils restent donc fermés sur eux-mêmes.

Demain, au contraire, les bibliothèques peuvent être des acteurs du web sémantique par les informations contenues dans leurs catalogues. Pour cela, la structure et les formats des catalogues doivent évoluer mais cela signifie que les règles de catalogage, elles aussi, évoluent. C'est tout l'enjeu du nouveau code de catalogage qu'est RDA !



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Questions de la salle

■ *Patrice Verrier*, médiathèque de la Cité de la Musique

A-t-on déjà des exemples en ligne de mise en œuvre de RDA, en dehors des quatre grandes bibliothèques nationales qui vont bientôt s'y mettre ?

■ *Françoise Leresche*

RDA est encore en cours d'élaboration. Le texte stabilisé est annoncé pour le 3^e trimestre 2009. Le 4^e trimestre devrait permettre d'affiner la formation, les scénarios d'implémentation dans les quatre bibliothèques nationales concernées et dans un groupe d'une vingtaine de bibliothèques qui représentent la diversité des établissements utilisateurs des AACR. L'IFLA organisera, lors de son congrès en 2010, un séminaire satellite qui sera consacré en grande partie à une première évaluation des expérimentations d'implémentation de RDA. Mais pour l'instant, je ne peux vous donner aucune adresse URL de bibliothèque qui aurait déjà commencé à l'implémenter.

■ *Elizabeth Giuliani*, adjointe à la directrice du département de l'Audiovisuel, BnF

Lorsqu'une opportunité de travailler sur ces normes se présente, il faut absolument la saisir. C'est un peu rude et complexe mais, en groupe, nous y arrivons mieux que tout seul. Cela permet non seulement d'apprendre mais aussi de faire profiter la communauté de ses réflexions et de ses expériences. L'interopérabilité est une chance de désenclavement des documents dits spécialisés qui deviennent plus facilement « RDAisables » que les imprimés classiques.

■ *Une personne du public*

Pour les analphabètes, dont je fais partie, pourriez-vous nous donner une petite explication sur FBR et FRAD ?

■ *Françoise Leresche*

Le modèle FRBR est un modèle entités-relations. Il définit les objets conceptuels manipulés lorsqu'on fait une description bibliographique (les entités). Les « entités » sont également les objets fondamentaux pertinents pour les utilisateurs de données bibliographiques.

Ces entités sont réparties en trois groupes :

- Un groupe 1, le plus novateur, décrit les quatre entités qui correspondent aux étapes de la production intellectuelle ou artistique :
 - l'**œuvre** qui est l'entité la plus abstraite ;



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- l'**expression** qui correspond aux moyens choisis pour exprimer l'œuvre (texte écrit ou parlé, musique notée ou interprétée, langue, etc.) ;
- la **manifestation** correspond à la publication éditoriale par un éditeur ;
- les **exemplaires** ou items sont les objets physiques qu'on manipule entre des œuvres, expressions, manifestations, documents.
- Un deuxième groupe correspond aux entités qui entretiennent des relations de responsabilité avec les différentes entités du groupe 1 :
 - **personne** ;
 - **collectivité** ;
 - **famille**.
- Un troisième groupe correspond aux relations de sujet que les entités peuvent entretenir vis-à-vis de l'œuvre. Il comprend aussi bien les entités du groupe 1 et du groupe 2 que quatre autres entités qui sont les **concepts**, les **événements**, les **lieux** et les **objets**.

Ces entités ont des attributs qui sont les grandes caractéristiques qui vont permettre de les décrire. Par exemple, pour une manifestation, entre autres attributs, figurent le titre propre, le nom de l'éditeur, la date d'édition, etc. Ensuite, sont définies les relations entre ces entités : par exemple, la relation entre une personne et une œuvre correspond à une mention de responsabilité qui donnera lieu à un point d'accès auteur.

Je vous incite à aller regarder la page de présentation de ce modèle sur le site web de la BnF, dans la rubrique *Informations pour les professionnels*. Vous y trouverez également une présentation des *Principes internationaux de catalogage*, etc. Le modèle FRBR est de plus en plus actuel et important à connaître car, désormais, on raisonnera à partir de ce modèle pour cataloguer.

■ *Une personne du public*

J'ai une autre question d'analphabète : j'ai beaucoup entendu parler du web sémantique depuis le début de l'après-midi. Pour moi, cette expression reste très abstraite...

■ *Françoise Leresche*

Le web traditionnel est plutôt un web de documents alors que le web sémantique serait plutôt un web de données qui sont exploitées et mises en relation.



Musique numérique : oubli dans l'âme ou supplément d'âme ?

■ *Elizabeth Giuliani*

Avec l'évolution des règles de catalogage, on parle de moins en moins de notices mais de données mises en relation entre elles, pour tisser un maillage d'informations qui sont moins encloses qu'une notice considérée comme un document.

■ *Françoise Leresche*

Passer au web sémantique, c'est permettre l'exploitation des données de nos catalogues et de nos notices pour d'autres usages et par d'autres utilisateurs du web.

Liens utiles

FRBR

Groupe de révision de l'IFLA : www.ifla.org/VII/s13/wgfrbr/

Texte anglais : www.ifla.org/VII/s13/frbr/

Traduction française : www.bnf.fr/pages/infopro/normes/pdf/FRBR.pdf

FRAD

Groupe de travail de l'IFLA : www.ifla.org/VII/d4/wg-franar.htm

Principes internationaux de catalogage

IME-ICC : www.ifla.org/VII/s13/icc/

Texte anglais : www.ifla.org/VII/s13/icp/ICP-2009_en.pdf

Traduction française : www.ifla.org/VII/s13/icp/ICP-2009_fr.pdf

Présentation sur le site de la BnF : www.bnf.fr/pages/infopro/normes/no-international.htm

RDA

Présentation générale : www.collectionscanada.ca/jsc/rda.html

Projet final soumis à enquête : www.rdaonline.org/constituencyreview/

Brochure en français : www.collectionscanada.gc.ca/jsc/docs/rdabrochure-fre-fr.pdf

Présentations dont celles du congrès satellite de l'IFLA à Québec :

www.collectionscanada.gc.ca/jsc/rdapresentations.html

Annnonce sur le site de l'IFLA : www.ifla.org/VII/s13/announcement/s13-Press24102007-fr.htm



Musique numérique : oubli dans l'âme ou supplément d'âme ?



3.4 LE SON DANS GALLICA⁹

par **PASCAL CORDEREIX**, chef du service des documents sonores, département de l'audiovisuel, BnF

Pour rebondir sur l'exposé de Jérôme Dupont et Thomas Ledoux à propos du SPAR, il faut bien prendre conscience que ce qui est diffusé comme document sonore dans Gallica connaît en amont une numérisation de conservation qui, elle, est effectuée dans une haute définition de 24 bits-96 kHz. Il faut donc considérer Gallica comme un produit dérivé de cette numérisation première. Ceci explique le choix d'un format de diffusion très « grand public » dans Gallica, du MP3 compressé à 192 kbs (320 kbs pour les livraisons à venir). Nous sommes dans du fichier « jetable », sachant qu'il existe toujours en amont la copie de conservation.

Les objectifs

La diffusion de documents sonores sur Gallica répond à deux objectifs complémentaires :

- rendre accessible au plus grand nombre un patrimoine sonore national et, notamment, au public distant en France ou à l'étranger ;
- rendre accessibles des répertoires peu ou pas réédités, sachant que les collections sonores du département audiovisuel couvrent maintenant plus d'un siècle d'histoire de l'enregistrement sonore puisque les fonds les plus anciens que nous conservons remontent à 1891.

Un service

La diffusion en ligne dans Gallica est un service qui a été lancé en janvier 2009. Il existait auparavant des documents sonores dans certains dossiers de Gallica, notamment dans *Voyages en France*, *Voyages en Afrique*, *La voix dans Gallica*, etc., mais il s'agissait de sélections parcellaires. Nous sommes passés dans une autre dimension, comme le soulignait Jacqueline Sanson ce matin, avec ce service encore dans sa phase expérimentale, mais avec de fortes ambitions à long terme.

Je signalerai rapidement trois nouveautés dans Gallica qui me paraissent intéressantes :

- la citabilité des ressources par copie du permalien ;
- la syndication par flux RSS ;

⁹ <http://gallica.bnf.fr>



Musique numérique : oubli dans l'âme ou supplément d'âme ?

- la création d'un espace personnel qui permet de constituer un panier de documents, de sauvegarder sa recherche, de la tagger pour regrouper un certain nombre de documents et reconstituer son propre espace documentaire.

À l'heure actuelle, notre corpus compte 1 000 documents (environ 2 000 titres et 80 heures d'enregistrement). Ce sont des documents libres de droit, en écoute et téléchargement gratuits.

La période potentiellement couverte est 1891-1958, sachant que nous n'avons pas encore livré les documents les plus anciens, les cylindres de Gustave Eiffel enregistrés en 1891.

Deux grandes composantes se distinguent d'ores et déjà dans ce corpus :

- des archives inédites ;
- des documents édités représentatifs de l'histoire de l'édition phonographique.

Les corpus

Quatre grands corpus sont représentatifs de ce fonds :

Pour les archives inédites :

- les Archives de la Parole (1911-1914) : la période fondatrice de notre département avec le linguiste Ferdinand Brunot ;

Pour les documents édités :

- le don fait par Émile Pathé aux Archives de la Parole en 1911 : il s'agit d'un fonds de 1 000 disques pour lequel nous avons retenu les 300 références qui correspondent environ à 600 enregistrements russes et caucasiens. C'est un fonds particulièrement rare qui se trouve uniquement à la BnF ;
- une sélection d'enregistrements de musiques baroque et ancienne, [ca] 1912-1958 ;
- une sélection d'enregistrements de musiques du monde, [ca] 1926-1949.

Les Archives de la Parole

Elles contiennent 2 composantes pour cette période comprise entre 1911 et 1914 :

- les voix « célèbres » de l'époque, enregistrées à la Sorbonne (Guillaume Apollinaire, Émile Durkheim, Maurice Barrès...),
- trois enquêtes de terrain menées entre 1912 et 1913, dans les Ardennes (1912), dans le Berry et le Limousin (1913). Il ne s'agit pas des premières enquêtes phonographiques faites en



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique : oubli dans l'âme ou supplément d'âme ?

France, mais ce sont les premières enquêtes aussi systématiques et constitutives d'un produit à grande échelle, en vue de constituer un atlas linguistique des parlers de France.

The screenshot shows a web browser window displaying a digital archive page. The main content is a circular image of a gramophone record with the text "UNIVERSITE DE PARIS ARCHIVES DE LA PAROLE" and "DISCOURS D'INAUGURATION DES ARCHIVES DE LA PAROLE M. BRUNOT". The interface includes a search bar, navigation controls, and a sidebar with a table of contents.

Ferdinand Brunot, *Discours d'inauguration des Archives de la Parole*, 3 juin 1911.

Pour écouter : <http://gallica.bnf.fr/ark:/12148/bpt6k1279113.r=archives+de+la+parole.langFR>

En parallèle de ce premier corpus de voix célèbres à la Sorbonne, Ferdinand Brunot a mené les trois enquêtes de terrain précitées. Par exemple, dans l'enquête pour le Berry, la *Briolée aux bœufs*, enregistrée en 1912, chant de travail et de conduite de troupeaux, témoigne d'une expression pratiquement disparue des campagnes.

Le don Pathé aux Archives de la parole en 1911

Il s'agit principalement d'enregistrements de musiques traditionnelles et de musiques populaires d'Europe de l'Est, russes et caucasiennes, représentatives de la succursale russe de Pathé, basée à Moscou, mais aussi de l'empire Pathé (avec ses succursales à Moscou, en Chine et à New York), tel qu'il a pu exister jusqu'à la fin de la Première Guerre mondiale.

Les deux autres corpus sont des sélections représentatives de l'histoire de l'édition phonographique :
Une sélection d'enregistrements de musiques baroque et ancienne, [ca] 1912-1955



Musique numérique : oubli dans l'âme ou supplément d'âme ?

Ce parti pris s'explique par l'existence d'une demande émanant de chercheurs pour travailler sur des répertoires difficilement accessibles, car quasiment pas réédités en dehors de quelques grandes figures de l'interprétation.

Il était également important pour nous de montrer ce qu'était la musique baroque, du point de vue de l'enregistrement sonore de l'édition phonographique, avant les « baroqueux », et de donner accès aux grandes marques phonographiques, comme l'Anthologie sonore, la Boîte à musique ou l'Oiseau-Lyre.

Une sélection d'enregistrements de musiques du monde, [ca] 1926-1949

Notre objectif était là aussi de donner accès à des enregistrements méconnus aujourd'hui, mais inscrits dans une filiation historique avec ce qui est produit depuis 30 ans en matière de *world music*. Il s'agit d'une sélection de 600 enregistrements environ, issus de la collection personnelle de Roger Dévigne, premier directeur de la Phonothèque nationale en 1938, folkloriste et amateur pointu de musiques du monde.

Les perspectives d'avenir

La mise en ligne des Archives de la Parole et du Musée de la Parole et du Geste sera poursuivie et, notamment, avec les enquêtes de terrain réalisées par Hubert Pernot et Roger Dévigne.

Après cette première offre, nous souhaiterions accroître celle de documents édités, représentatifs de l'histoire de l'édition phonographique, comme les cylindres et les 78 tours.

Nous sommes encore dans les prémices d'un projet amené à se développer.

Les 1 000 premiers documents mis en ligne étant des documents libres de droits, qu'en sera-t-il de l'offre de documents sonores sous droits dans Gallica ? Cette question est actuellement en étude au sein de la BnF.

Nous avons également comme perspective la constitution d'une plateforme nationale de téléchargement de disques 78 tours.



ACIM

Bulletin de liaison électronique n°07 – septembre 2009

Musique numérique :
oubli dans l'âme ou supplément d'âme ?

Retrouvez les bulletins précédents sur le portail des bibliothécaires musicaux :

<http://www.acim.asso.fr/spip.php?rubrique90>

Si vous ne souhaitez plus recevoir cette lettre ou si vous rencontrez un problème de réception, merci de nous en avvertir par mél. :

lemaire.acim@gmail.com

ACIM (Association pour la Coopération des professionnels de l'Information Musicale)

Siège social :

46bis rue Saint-Maur – 75011 Paris

SIRET : 382.220.945.00069 – Code APE : 9101Z

Adresse pour la trésorerie et les adhésions :

c/o Patrick Goczkowski

14 avenue des tilleuls – 95320 Saint-Leu-la-Forêt

Bulletin d'adhésion sur :

<http://acim.asso.fr/spip.php?article213>

