

Journée réseau du centre régional du Sudoc-PS sciences juridiques, politiques et économiques (CR 27), hébergée à la Bibliothèque de l'Hôtel de Ville de Paris

28 février 2013

La gestion des ressources continues à l'heure de la numérisation



Visite organisée de la Bibliothèque de l'Hôtel de Ville pour le réseau du CR 27 (© BHDV)

Mot d'accueil par M. Pierre Casselle, directeur de la Bibliothèque de l'Hôtel de Ville

I – Le catalogue Sudoc : point d'entrée de la gestion des collections du CR 27¹

L'après-midi s'est ouvert sur un bilan de l'activité du CR 27 depuis la dernière réunion du réseau en mai 2011. En un peu plus de 20 mois, le CR a signalé 2.213 états de collections supplémentaires ce qui nous mène, au total, à 22.798 localisations. Cela a donné lieu à une amélioration qualitative du catalogue (créations, modifications et suppressions de notices). De plus, certains établissements ont commencé à constituer des dossiers issn, instaurant une dynamique que le CR 27 souhaiterait voir se développer.

Le nombre d'Unica s'élève à 1.674 et la liste de ces titres a été donnée à chacune des bibliothèques qui en possèdent. Il s'agit d'exemplaires uniques en France. Il n'est malheureusement pas possible d'obtenir les Unica à l'échelle de l'Ile-de-France, ce qui constituerait aussi une précieuse indication. Pour pallier ce manque, il est possible, *via* l'outil Periscope ou directement au titre de la notice dans le Sudoc, de repérer les autres localisations.

¹ PPT de l'intervention : http://bcujas-cms.univ-paris1.fr/sites/default/files/documents/Activites_du_CR_27_juin_2011_-_janvier_2013.pdf

Du fait du travail effectué depuis 4 ans, le Sudoc représente désormais un outil indispensable pour gérer ses collections (dons, pilons, suspensions de titres etc...) et, à moyen ou long terme, pour établir une carte documentaire pour le CR 27.

Les conventions triennales de partenariat sont en cours de renouvellement. Pour 2014, sont concernés : l'ENA, le Conseil d'Etat, l'Unesco, le CRDALN, la Cour des Comptes, la DRASS, l'IFRI.

L'Abes va proposer d'ici le printemps 2013 un nouvel outil : **Colodus**, pour permettre aux bibliothèques du Sudoc de saisir leurs états de collections à partir d'une application web et non plus de WinIBW. Le CR 27 devrait proposer ce nouveau service aux bibliothèques du CR à partir de la rentrée 2013. En premier lieu, les bibliothèques dites "entrées secondaires" seront contactées, puis les autres bibliothèques du réseau qui souhaiteraient en bénéficier.

II – Les plans de numérisation dans le CR27

Les exemples de Cujas (Alexandra Gottely) et du CEDIAS (Anthony Lorry).

1) La numérisation à Cujas²

• **Les balbutiements**

Le 1er essai date de 2001 et ne concernait que 17 titres, numérisés par la société Safig. Les images étaient conservées sur les ordinateurs de la société et ont été détruites par un virus. Le peu qui a été récupéré n'est pas exploitable.

Une 2e expérience a été réalisée en 2004 dans le cadre d'une exposition virtuelle pour le bicentenaire du code civil. Elle n'a concerné que quelques ouvrages (cf. <http://expocujas.univ-paris1.fr/Bicentenaire/default.htm>)

En 2007, la bibliothèque achète un scanner (prototype) et teste la numérisation sur les cours photocopiés du doyen Carbonnier. La numérisation a été effectuée en interne et suivie d'une océrisation entièrement relue et corrigée par 8 personnes pendant 4 mois, pour un total de 3.000 pages. Cette expérience, jugée fastidieuse et chronophage, ne sera pas renouvelée.

• **Actuellement**

La bibliothèque dispose de plusieurs programmes de numérisation :

Tout d'abord un programme de numérisation en interne et en externe. Les documents numérisés en interne sont postérieurs à 1830 (le scanner ne supporte pas les documents trop anciens) et les documents plus anciens sont numérisés à l'extérieur par un prestataire (Arkhenum).

Ensuite, une collaboration ANR a été mise en place avec l'Institut Michel Villey. 4.500 pages ont été numérisées, océrisées, puis relues de manière exhaustive par une vacataire recrutée uniquement pour cette tâche. Cela a abouti à la création du site "Encyclopédie de droit politique" : <http://www.droitpolitique.com/spip.php?page=sommaire-edp>.

² PPT de l'intervention : http://bcujas-cms.univ-paris1.fr/sites/default/files/documents/La_numerisation_a_Cujas_CR27_A_Gottely.pdf

Une autre collaboration ANR a été mise en place avec le laboratoire CECOJI pour le projet “Mémoires” qui prévoit de numériser toutes les grandes lois du patrimoine. Pour l'instant, ce projet est en attente.

Les reliés Cujas. Il s'agit des documents envoyés à la BnF pour être numérisés. En 2010, 43 titres de la liste Pfister-Roumy ainsi que le Bulletin de la Cour de Cassation (de 1798 à 1909, chambres civiles et criminelle) sont passés sur les chaînes de numérisation de la BnF. En 2012, ce fut le tour des journaux officiels des anciennes colonies. Ce processus est assez lourd pour la bibliothèque qui doit rédiger des “fiches-état” en ligne, mettre les documents en caisses et gérer leurs allers retours entre les deux établissements. En 2 mois et demi de travail, une personne à temps plein gère 240.000 pages.

Enfin, la Bibliothèque nationale de France et Cujas copilote un programme de numérisation et de valorisation concertées en sciences juridiques. Cette initiative réunit les grands corps de l'Etat, des professionnels du droit (enseignants, chercheurs et praticiens) et des bibliothèques de tout statut, autour de la numérisation de corpus juridiques définis en concertation. Il s'agit d'un appel à initiative concernant des projets de numérisation sur des corpus moyens (projets sur 1 ou 2 ans). La BnF subventionne entre 50 et 70 % des frais de numérisation.

- **Concrètement**

Le choix des titres à numériser : il se fait grâce à la liste Pfister-Roumy (2 professeurs en histoire du droit) établie en 2007 et qui comprend 436 titres majeurs (datant du 15^e au 20^e siècle) soit 1.400 volumes, 893.000 pages. Pour consulter cette liste : <http://cujasweb.univ-paris1.fr/search#>

La création de lots : pour la numérisation externe, les ouvrages ont été répartis en lots, en fonction de critères uniquement physiques.

Les tests : Les images du prestataire sont contrôlées par la bibliothèque Cujas par sondage, qui valide ou non la qualité du train de numérisation. Pour la numérisation en interne, le contrôle des images se fait en temps réel, à chaque image.

La table des matières et le bordereau d'état : chaque document est catalogué et indexé (une indexation particulière est nécessaire pour l'interrogation par facettes dans la bibliothèque numérique de Cujas). La table des matières permet la navigation dans le document et sert de table de concordance. Elle est hiérarchisée et réalisée par des vacataires sur un fichier Excel. Le bordereau d'état est réalisé pour le prestataire extérieur : page par page, on note les défauts, l'état physique du document, etc. La 1^{ère} partie du bordereau concerne la notice bibliographique telle quelle doit être reprise par le prestataire dans la notice Dublin core.

Après la numérisation : en interne, on procède à une relecture de l'OCR pour les points d'entrée dans le document uniquement (titres, tables des matières, index). Puis on rédige la notice Dublin core (en langage XML) qui correspond à la notice du livre dans la bibliothèque numérique. Elle reprend les éléments de la notice ISBD. Juste avant la mise en ligne, on intègre sous forme de signets dans le pdf, la table des matières de chaque document et on met en place sur chaque image un filigrane “BIU Cujas”. Le moissonnage est réalisé grâce à un entrepôt OAI (qui peut être moissonné par Gallica). L'archivage pérenne se fait avec le CINES.

L'espérance de vie des documents numérisés actuellement est de 1 à 5 ans. Pour agrandir cette durée de vie, 2 stratégies sont possibles : la migration (régulièrement on change le format des fichiers en remplaçant l'ancien type de fichier par un plus récent. Ex. : .doc et .docx pour les documents Word) et l'émulation (on lit les fichiers sur des ordinateurs disposant d'un logiciel qui trompe l'ordinateur en lui faisant croire qu'il est sous un ancien système). En France, 2 grands organismes gèrent l'archivage pérenne. Il s'agit

du CINES pour l'enseignement supérieur (5.000€/téraoctet) et de la BnF pour tous les autres (2.000€/téraoctet). Un document numérisé représente au moins 20h de travail.

2) La numérisation au CEDIAS³

• **Présentation :** www.cedias.org/

Le CEDIAS est une fondation privée reconnue d'utilité publique créée en 1894. Elle couvre de nombreux domaines de connaissances dont : l'économie sociale, la mutualité, le syndicalisme, l'hygiène, le logement social, etc.

Sa bibliothèque compte plus de 120.000 volumes (dont 70.000 monographies) soit 3 km linéaires. L'établissement est soumis à de nombreuses contraintes : conservation des documents, valorisation des collections, peu de personnel (2 titulaires et parfois 1 aide vacataire) et aucun budget d'acquisition → appel aux dons réguliers.

En 2000, le virage du numérique a été abordé par le biais de la ré-informatisation du catalogue *via* le logiciel libre PMB. Le SIGB a été conçu pour fédérer plusieurs types de bases. Le parti pris de la bibliothèque du CEDIAS était de n'avoir qu'une seule application pour toutes les ressources documentaires (archives, images, documents numériques, catalogue). Le logiciel PMB, sans cesse en évolution, permet par exemple d'utiliser une visionneuse pour la photothèque.

Le CEDIAS est ensuite devenu pôle associé de la BnF en 2008 et un programme de rétroconversion a été entrepris de 2009 à 2011. Enfin, le classement de la bibliothèque en archives historiques a permis de préserver l'espace documentaire (bibliothèque et magasins).

• **La politique de numérisation**

2 types de partenariats avec la BnF ont été mis en place : Les **reliés extérieurs** (les collections sont confiées à la BnF pour être numérisées sur sa chaîne de numérisation), dès 2009, et la coopération (la BnF cofinance à hauteur de 50% la prestation de numérisation à condition de stocker les données dans un entrepôt OAI pour que Gallica puisse les moissonner) qui nécessite toutefois, pour le CEDIAS, le recours au mécénat pour trouver les 50% de la somme non prise en charge par la BnF. Par les reliés extérieurs, le CEDIAS numérise 90.000 pages 2 fois par an.

• **La numérisation BnF**

Il faut compter 6 mois entre le départ du train du CEDIAS et la mise en ligne totale dans Gallica. Les documents sont personnalisés dans Gallica par un logo de la bibliothèque partenaire sur chaque document. Le CEDIAS a choisi de numériser des documents encore soumis au droit d'auteur ce qui entraîne des démarches auprès des auteurs et ayants droits pour obtenir le droit de numériser et diffuser ces documents.

Normalement la BnF doit fournir une copie des fichiers numérisés. Elle le fait mais dans un format qui lui est propre (en fait un ensemble de 5 formats informatiques différents) et qui rend donc les fichiers difficilement exploitables par les bibliothèques partenaires.

³ PPT de l'intervention : <http://bcujas-cms.univ-paris1.fr/sites/default/files/documents/CEDIAS-CR27-2013-1.pdf>

PMB a réussi à mettre au point une “moulinette” informatique (la dernière version en développement contient une visionneuse Open Library) pour transformer les fichiers de la BnF en fichiers PDF récupéré au sein du catalogue.

- **Bilan à l'horizon 2013**

300.000 pages seront numérisées, dont plusieurs revues sous droits. Des collections complètes ont été traitées grâce à un travail en réseau (associations professionnelles, réseaux documentaires, etc.) qui a permis de combler les lacunes des collections du CEDIAS et ainsi proposer des collections complètes.

- **Et après**

Volonté de valoriser les collections du CEDIAS dans l'optique d'accueillir un public très spécialisé.

La BnF numérise maintenant en couleur ce qui crée des problèmes car les fichiers sont très lourds à charger. *Internet Archive* propose depuis 1 an une visionneuse (HTML5) qui permet de charger plus rapidement les fichiers lourds et compatible avec les produits Apple puisque n'utilisant pas le Flashplayer. Cette visionneuse permet de “pré-générer” un fichier PDF pour que le lecteur puisse le charger plus rapidement. Le CEDIAS compte utiliser cette visionneuse à partir de 2013.

- **L'entrepôt OAI du CEDIAS** : <http://cediasbibli.org/page/entrepot-oai>

Pour faciliter l'accès à ses collections numérisées, la bibliothèque du CEDIAS a mis en place un entrepôt OAI-PMH. Les notices sont disponibles en format Dublin Core, et ont été paramétrées suivant les recommandations de la Bibliothèque nationale afin d'être indexées dans Gallica. L'entrepôt du CEDIAS-Musée social est moissonné par plusieurs sites : Gallica, ISIDORE, plateforme de recherche permettant l'accès aux données numériques des sciences humaines et sociales (SHS) et [BASE](#) (Bielefeld Academic Search Engine) : l'un des moteurs de recherche les plus volumineux du monde spécialisé pour des données numériques universitaires disponibles en accès libre.

III- Un nouvel outil de gestion des collections : Périscope.

Cf. http://bcujas-cms.univ-paris1.fr/sites/default/files/documents/Periscope_Drupal_mars_2013.pdf

IV – Visite de la Bibliothèque de l'Hôtel de Ville



Merci aux collègues de la BHDV pour leur accueil, la préparation d'une exposition des documents rares et précieux et l'organisation des deux groupes de visite.

La Bibliothèque de l'Hôtel de Ville a été constituée à partir de 1872 pour remplacer celle détruite dans l'incendie du 24 mai 1871. Elle ouvre à nouveau ses portes le 15 juin 1890 et compte aujourd'hui près de 600 000 volumes, 7 000 titres de périodiques, 2 270 manuscrits et des fonds importants de dessins d'architecture (fonds Gabriel Davioud, Théodore Ballu et Édouard Deperthes, Joseph Bouvard...) ou de photographies (Charles Marville, construction du métropolitain et de divers équipements municipaux, portraits...).

Les fonds eurent dès l'origine un caractère quasi encyclopédique dans le souci de couvrir tous les centres d'intérêt de l'administration parisienne. A noter, les fonds étrangers - l'une des ressources les plus originales de la bibliothèque - sont le fruit des échanges internationaux organisés par la Ville de Paris à l'initiative d'Alexandre Vattemare, du milieu du XIX^e siècle aux années 1920. Une politique de numérisation a été engagée depuis plusieurs années en collaboration avec la Parisienne de Photographie (5 000 images) et la bibliothèque est également associée, pour sa collection de publications administratives parisiennes, au programme de numérisation concertée en sciences juridiques conduit par la Bibliothèque nationale de France.

La salle de lecture, d'une superficie de 600 m² a été conçue par l'architecte Édouard Deperthes et l'ensemble fut entièrement rénové en 2010.