

Diplôme de conservateur de bibliothèque

Mémoire d'études / janvier 2015

L'ABES et la BnF en route vers l'*open data*

Clémence Agostini

Sous la direction de Gildas Illien
Directeur du département de l'Information Bibliographique et Numérique
de la BnF

Remerciements

Je tiens à remercier chaleureusement toutes les personnes qui ont accepté de me rencontrer dans le cadre de ce mémoire et qui ont consacré du temps à me parler de leur métier, avec beaucoup de bienveillance et d'amabilité.

Je remercie tout particulièrement Gildas Illien pour avoir proposé ce sujet passionnant et pour m'avoir guidée tout au long de cette étude.

Enfin merci à mes proches, à Karine pour ses conseils avisés, à Katell et Adèle pour leur relecture et leur intérêt, ainsi qu'à Line et Jean-Claude.

Résumé :

L'ABES et la BnF, les deux agences bibliographiques françaises, se sont engagées dans le mouvement d'ouverture des données publiques en plaçant leurs métadonnées bibliographiques sous la licence ouverte de l'État et en adoptant les standards du web sémantique. Ce mouvement accompagne une mutation dans laquelle sont engagées les bibliothèques et qui vise à faire évoluer leurs dispositifs de signalement, ceux-ci étant devenus inadaptés aux ressources électroniques et aux nouveaux usages de recherche d'information sur le web. Cette « transition bibliographique » va provoquer des changements importants dans les modes de travail au sein des bibliothèques. Pour faire face à ces mutations, l'ABES et la BnF, deux institutions qui diffèrent par leurs missions et leurs modalités de fonctionnement mais qui font face aux mêmes enjeux, ont mis en œuvre une démarche de coopération.

Descripteurs :

Administration publique et Internet, Bibliothèques et Internet, Catalogage, Web sémantique

Abstract :

The bibliographic agency of academic libraries (ABES) and the French national library (BnF) have committed to the open data movement, putting their bibliographic metadata under the French State Open Licence and adopting linked open data standards. This movement goes with a mutation known as “bibliographic transition” consisting in adapting cataloging systems to electronic resources and to the evolution of information search on the web. This transition is going to make important changes happen in the library profession. To face these mutations, ABES and BnF, two institutions with different missions but facing similar stakes, have decided to cooperate.

Keywords :

Internet in public administration, Libraries and the Internet, Cataloging, Semantic Web

Droits d'auteurs



Cette création est mise à disposition selon le Contrat :

Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France

disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/2.0/fr/> ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.

Sommaire

SIGLES ET ABRÉVIATIONS.....	7
INTRODUCTION.....	9
PRÉAMBULE : CONTEXTE ET DÉFINITIONS.....	11
LES INSTITUTIONS EN PRÉSENCE.....	18
Deux agences bibliographiques.....	18
La Bibliothèque nationale de France.....	19
<i>Missions.....</i>	<i>20</i>
<i>Le réseau de la BnF.....</i>	<i>20</i>
<i>L'agence bibliographique nationale.....</i>	<i>23</i>
<i>Actualité de la BnF.....</i>	<i>24</i>
L'Agence Bibliographique de l'Enseignement Supérieur.....	27
<i>La naissance de l'ABES.....</i>	<i>27</i>
<i>La diversification des activités de l'ABES.....</i>	<i>29</i>
<i>Les nouveaux projets.....</i>	<i>32</i>
Les débuts de la coopération.....	35
<i>Le signalement des ressources.....</i>	<i>35</i>
<i>La mise en commun d'une réflexion stratégique.....</i>	<i>36</i>
<i>Des rôles bien définis.....</i>	<i>37</i>
LA TRANSITION BIBLIOGRAPHIQUE.....	39
L'évolution des collections, des besoins et des usages.....	39
<i>Les ressources numériques changent la donne.....</i>	<i>39</i>
<i>Le web change la donne.....</i>	<i>40</i>
<i>Bibliothèques et web de données.....</i>	<i>42</i>
Vers un nouveau modèle pour la description bibliographique.....	45
<i>Nos outils actuels.....</i>	<i>45</i>
<i>Une nouvelle façon de modéliser l'information : FRBR.....</i>	<i>46</i>
<i>Un nouveau code de catalogage : RDA.....</i>	<i>49</i>
<i>Vers une révision des formats MARC.....</i>	<i>51</i>
Vers de nouveaux modes de travail ?.....	53
<i>Du catalogage au « cataliage ».....</i>	<i>53</i>
<i>Une nouvelle répartition du travail entre les acteurs.....</i>	<i>56</i>
OUVRIER LES DONNÉES : LES STRATÉGIES MISES EN ŒUVRE.....	59
S'insérer dans le web de données.....	60
<i>L'approche de l'ABES.....</i>	<i>61</i>
<i>L'approche de la BnF.....</i>	<i>63</i>
<i>Bilan.....</i>	<i>65</i>
<i>Interrogation croisée des données.....</i>	<i>66</i>
Mutualiser les autorités.....	66
<i>Les personnes et collectivités.....</i>	<i>67</i>
<i>Les sujets.....</i>	<i>71</i>
<i>Les titres et œuvres.....</i>	<i>72</i>
<i>Vers un fichier national d'autorités ?.....</i>	<i>74</i>
Préparer la transition bibliographique.....	74
<i>Un nouveau cadre pour adopter RDA.....</i>	<i>74</i>
<i>Conduire le changement.....</i>	<i>76</i>
CONCLUSION.....	82

SOURCES.....	85
BIBLIOGRAPHIE.....	86
Institutions.....	86
Open data.....	86
Métier.....	87
Signalement.....	88
Web, web sémantique, bibliothèques et web de données.....	91
TABLE DES MATIÈRES.....	93

Sigles et abréviations

AACR : *Anglo American Cataloguing Rules*

ABES : Agence bibliographique de l'enseignement supérieur

AERES : Agence d'évaluation de la recherche et de l'enseignement supérieur

AFNOR : Association française de normalisation

API : *Application Programming Interface*

BIBFRAME : *Bibliographic Framework Initiative*

BnF : Bibliothèque nationale de France

BSN : Bibliothèque scientifique numérique

CADA : Commission d'accès aux documents administratifs

CBU : Contrôle bibliographique universel

CCFr : Catalogue collectif de France

CENL : *Conference of European National Librarians*

CNCPB : Centre national de coopération des bibliothèques publiques

CNRS : Centre national de la recherche scientifique

CSB : Comité stratégique bibliographique

DNB : *Deutsche Nationalbibliothek*

EURIG : *EUropean RDA Interest Group*

FOAF : *Friend Of A Friend*

FRAD : *Functional Requirements for Authority Data*

FRBR : *Functional Requirements for Bibliographic Records*

FRSAD : *Functional Requirements for Subject Authority Data*

HTML : *Hypertext Markup Language*

HTTP : *Hypertext Transfer Protocol*

IBN : Département de l'information bibliographique et numérique (BnF)

IFLA : *International Federation of Library Associations and Institutions*

IGB : Inspection générale des bibliothèques

INIST : Institut de l'information scientifique et technique

ISBD : *International Standard Bibliographic Description*

ISNI : *International Standard Name Identifier*

ISO : *International Organization for Standardization*

JSC : *Joint Steering Committee for Development of RDA*

LCSH : *Library of Congress Subject Headings*

LRU (loi) : loi du 10 août 2007 relative aux libertés et responsabilités des universités

MARC : *Machine Readable Cataloguing*
OCLC : *Online Computer Library Center*
ORCID : *Open Researcher and Contributor Identifier*
RAMEAU : Répertoire d'autorité-matière encyclopédique et alphabétique unifié
RDA : *Resources Description and Access*
RDF : *Resources Description Framework*
SCD : Service commun de la documentation
SGBM : Système de gestion de bibliothèque mutualisé
SIGB : Système intégré de gestion de bibliothèque
SKOS : *Simple Knowledge Organisation System*
SPARQL : *SPARQL Protocol and RDF Query Language*
SUDOC : Système universitaire de documentation
URI : *Uniform Resource Identifier*
VIAF : *Virtual International Authority File*
W3C : *World Wide Web consortium*
XML : *eXtensible Markup Language*

INTRODUCTION

En 2015, le gouvernement français est censé proposer une loi adaptant la directive européenne du 26 juin 2013¹. Celle-ci élargit l'ouverture des données publiques aux bibliothèques, musées et archives, signant en principe la fin de l'exception pour les données culturelles présente à l'origine dans la loi CADA².

Mais les bibliothèques n'ont pas attendu l'évolution de la législation pour s'impliquer dans le mouvement d'ouverture des données publiques (*open data*). L'Agence bibliographique de l'enseignement supérieur (ABES) et la Bibliothèque nationale de France (BnF) ont placé leurs métadonnées³ sous la licence ouverte de l'État respectivement le 23 novembre 2012 et le 1^{er} janvier 2014.

Certes, cette démarche s'inscrit dans le cadre des injonctions des pouvoirs publics et elle est conforme à l'air du temps. Mais elle accompagne aussi une mutation que connaissent actuellement les bibliothèques et à laquelle on se réfère parfois sous le nom de « transition bibliographique ». Cette mutation est intrinsèquement liée à l'apparition des ressources électroniques et à celle du web.

En effet, les bibliothèques ont fait le constat que leurs outils de signalement, conçus dans un monde sans Internet et pour des ressources essentiellement imprimées, n'étaient plus adaptés. Les bibliothèques se sont organisées très tôt sur le plan de l'informatisation, de la structuration et de l'échange de leurs données, mais essentiellement entre elles, dans un monde resté jusqu'à présent très cloisonné. Elles ont également pris conscience que leur offre documentaire devenait invisible, noyée dans la masse d'informations disponibles sur le web. L'ouverture des données entre donc dans cette stratégie d'adaptation des bibliothèques à leur environnement numérique, afin de devenir plus visibles et d'offrir des services adaptés aux nouveaux usages du web.

Les bibliothèques sont également confrontées à une masse de ressources à référencer qui augmente sans cesse et qui devient difficile à absorber, tout en évoluant dans un contexte budgétaire peu favorable. Pour faire face à cette situation, elles s'emploient à réviser en profondeur leurs processus de production et de diffusion de métadonnées bibliographiques. Cela se manifeste par une tendance à chercher à obtenir les données au plus près de leur source de production, afin d'économiser des charges de catalogage, ou par la mise en œuvre de stratégies de mutualisation. L'idée forte qui sous-tend cette mutation consiste en effet à lier ses données à celles produites par d'autres acteurs, y compris en-dehors du monde des bibliothèques.

La transition bibliographique se caractérise aussi par la volonté de changer les outils de travail actuellement utilisés. C'est dans ce cadre que s'inscrivent les réflexions autour de l'adoption du modèle FRBR (*Functional Requirements for Bibliographic Records*) et du code de catalogage RDA (*Resources Description and*

¹ Directive 2013/37/UE du Parlement européen et du Conseil du 26 juin 2013 modifiant la directive 2003/98/CE concernant la réutilisation des informations du secteur public. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:FR:PDF>

² Loi n° 78-753 dite loi CADA du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal. <http://legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000339241>

³ Notons que les données dont il est question dans cette étude sont des métadonnées (elles seront définies plus loin) et ne comprennent pas les documents des collections.

Access) ou sur le remplacement des formats MARC (*Machine Readable Cataloguing*).

Ces changements ne manqueront pas d'avoir des conséquences sur les façons de travailler. À la fois parce que la mutualisation des tâches de signalement implique une distribution du travail entre les différents acteurs qui nécessite de se demander qui fait quoi et selon quels critères. Cette question se pose à l'échelle nationale et internationale, car les bibliothèques évoluent dans un contexte mondialisé. Mais des répercussions se feront aussi sentir sur le plan de l'identité professionnelle, notamment avec la question du catalogage, qui évolue peu à peu de la production de données vers du traitement et de la création de liens entre des données existantes. Quelles conséquences cela aura-t-il sur les métiers des bibliothèques ?

La BnF et l'ABES, têtes de pont respectivement pour les bibliothèques publiques et celles de l'enseignement supérieur, se sont emparé de ces réflexions. S'il s'agit de deux agences bibliographiques, ce sont deux institutions bien différentes, par leurs missions, leur périmètre, leur culture. Elles sont néanmoins confrontées à des enjeux similaires : redonner de la visibilité à leur données et faire face à une augmentation des ressources à référencer sans augmentation de leurs moyens. Dans cette optique, et compte tenu de leur convergence d'intérêts, elles ont engagé une démarche de coopération. Celle-ci se manifeste par exemple dans la création en 2012 du Comité stratégique bibliographique, dans le but d'élaborer une vision stratégique commune sur toutes les questions de signalement.

Comment les deux agences bibliographiques françaises s'insèrent-elles dans le mouvement d'ouverture des données publiques ? Comment ces deux institutions qui diffèrent par leurs missions et leurs modalités de fonctionnement, mais qui font face à des enjeux comparables, peuvent-elles trouver des axes concrets de coopération ? A l'heure où la recherche d'efficacité s'impose, cette démarche répond-elle bien aux attentes de mutualisation que l'on est en droit d'avoir aujourd'hui ? Par ailleurs, comment ces nouvelles stratégies consistant à ouvrir et lier les données vont-elles affecter les modes de travail dans les établissements que les deux agences doivent servir ?

Un certain nombre d'éléments de contexte seront au préalable évoqués en préambule (définition de l'*open data*, contexte français) et les principaux concepts nécessaires à la compréhension du sujet seront définis (données produites par les bibliothèques, ouverture juridique, ouverture technique).

Puis nous examinerons comment la BnF et l'ABES, à travers leur histoire, leurs missions, leur actualité (I) se positionnent respectivement dans une dynamique d'adaptation aux enjeux de la transition bibliographique (II). Nous nous demanderons enfin quelles stratégies d'ouverture de leurs données elles adoptent et dans quelle mesure celles-ci leur permettent de mutualiser leur action (III).

PRÉAMBULE : CONTEXTE ET DÉFINITIONS

Qu'est-ce que l'open data ?

Le principe de l'ouverture des données publiques trouve en France son fondement dans la Déclaration des Droits de l'Homme et du Citoyen selon laquelle « la société est en droit de demander compte à tout agent public de son administration⁴ ». La loi du 17 juillet 1978 relative au droit d'accès aux documents administratifs (loi CADA) reconnaît à toute personne le droit de consulter « les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission⁵ ».

La circulaire du 14 février 1994 sur la diffusion des données publiques précise la notion de donnée, qui est « une information collectée ou produite sur n'importe quel support⁶ ». Une donnée publique est « une donnée collectée ou produite, dans le cadre de sa mission, par un service public, sur fonds publics⁷ ».

L'*open data* revêt selon Simon Chignard trois dimensions⁸. Il s'agit tout d'abord d'une caractéristique applicable aux données, qui peuvent être ouvertes selon trois critères, technique, juridique ou économique. Le critère économique signifie qu'il n'y a pas de redevance liée à la réutilisation d'une donnée, ou bien qu'elle est limitée à son coût marginal. Les critères juridiques et économiques seront approfondis plus loin. L'*open data* désigne par ailleurs un mouvement, mais aussi une injonction, une demande faite en direction des pouvoirs publics.

L'un des premiers exemples de l'appropriation de ce mouvement par les politiques publiques a lieu aux États-Unis, lors de la campagne électorale de Barack Obama en 2008. La volonté d'impliquer les citoyens dans la campagne s'est manifestée par la création du réseau social *mybarrackobama.com*. Par ailleurs Barack Obama a publié dès le premier jour de son mandat le Mémorandum « Transparence et gouvernement ouvert », dans lequel il manifeste sa volonté de faire appliquer la transparence, la participation et la collaboration⁹.

La Grande-Bretagne est également un pays précurseur pour les politiques d'*open data*. Dès 2006 deux journalistes du *Guardian* avaient lancé la revendication « *give us back our crown jewels*¹⁰ » fondée sur l'idée que dans la mesure où ce sont les impôts qui financent la production des données publiques, celles-ci doivent être disponibles pour les citoyens. La nomination de Tim Berners-Lee, figure emblématique, fondateur du web et dirigeant du W3C¹¹ au poste de conseiller pour l'*open data* a été un geste fort, suivi de la création du portail *data.gov.uk* en 2009. L'*open data* s'inscrit parfaitement dans la culture britannique de l'*accountability* (fait de rendre des comptes ou

⁴ Déclaration des Droits de l'Homme et du Citoyen de 1789. <http://www.legifrance.gouv.fr/Droit-francais/Constitution/Declaration-des-Droits-de-l-Homme-et-du-Citoyen-de-1789>

⁵ Loi n° 78-753 dite loi CADA du 17 juillet 1978 portant diverses mesures d'amélioration des relations entre l'administration et le public et diverses dispositions d'ordre administratif, social et fiscal. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000339241>

⁶ Circulaire du 14 février 1994 relative à la diffusion des données publiques. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000181007>

⁷ *ibidem*

⁸ CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*. Limoges, FYP éditions, 2012.

⁹ http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment

¹⁰ « Rendez-nous les joyaux de la couronne »

¹¹ Le W3C (*World Wide Web consortium*) est l'organisme de normalisation du web. <http://www.w3.org/>

transparence). Le portail *Where does my money go?*¹² montre ainsi la répartition des impôts dans le budget de l'État et permet à chacun en fonction de son salaire de connaître la part de ses impôts attribuée à la culture, à la sécurité, à la santé, à l'environnement...

Le contexte français

En France, comme l'indique Aurore Cartier dans son mémoire d'études¹³, l'*open data* prend sa source à l'échelon local. Des municipalités pionnières comme Rennes ou Paris se lancent dès 2010 dans l'ouverture de leurs données.

Au niveau du gouvernement, la mission Etalab est créée par décret¹⁴ le 21 février 2011 et rattachée depuis le 30 octobre 2012 au Secrétariat général pour la modernisation de l'action publique (SGMAP). Elle joue un rôle d'accompagnement des administrations pour l'ouverture de leurs données et administre le portail *data.gouv.fr*¹⁵, lancé le 5 décembre 2011. Une nouvelle version en est proposée depuis le 18 décembre 2013, plus ergonomique et avec un fonctionnement collaboratif.

Sur le plan européen, la directive 2013/37/UE adoptée le 26 juin 2013 par le Parlement européen et le Conseil de l'Union européenne modifie la directive dite PSI concernant la réutilisation des informations du secteur public¹⁶. Outre la fin de l'exception pour les données culturelles, cette directive prévoit la limitation des redevances d'utilisation aux coûts marginaux, et le recours à des critères de calcul de ces redevances qui soient objectifs et transparents. La directive encourage également l'utilisation de licences n'imposant pas d'autre restriction que la mention de la source, affichant ainsi la volonté que les licences ouvertes deviennent une pratique courante dans l'ensemble des pays de l'UE. Elle promeut également la mise à disposition des documents dans des formats ouverts et lisibles par des machines, interopérables, accompagnés de métadonnées les plus fines et précises possibles.

Le 18 juin 2013 les États membres du G8 ont adopté la *Charte du G8 pour l'ouverture des données publiques*¹⁷ qui affirme que les données publiques doivent être ouvertes par défaut, qu'elles doivent être de qualité et disponibles en quantité, accessibles et réutilisables par tous. Cette charte est déclinée dans chaque État membre par un plan d'action national, publié en France le 6 novembre 2013¹⁸. Il réaffirme ces grands principes et propose des axes de travail spécifiques pour les catégories de données identifiées comme ayant un fort potentiel. Le plan prévoit ainsi d'ici la fin de l'année 2015 la mise à disposition des données publiques sur les loyers.

Sur de nombreux points, cette charte et le plan d'action qui en découle ne font qu'entériner des avancées déjà existantes en France, pays relativement en avance et

¹² <http://wheredoesmymoneygo.org/>

¹³ CARTIER Aurore, *Bibliothèques et Open data, Et si on ouvrait les bibliothèques sur l'avenir ?*, mémoire DCB sous la direction de Frédérique Schlosser, 2013. Disponible sur : <http://www.enssib.fr/bibliotheque-numerique/documents/60401-bibliotheque-et-open-data-et-si-on-ouvrait-les-bibliotheques-sur-l-avenir.pdf>

¹⁴ Décret n° 2011-194 du 21 février 2011 portant création d'une mission « Etalab » chargée de la création d'un portail unique interministériel des données publiques. <http://legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000023619063&categorieLien=id>

¹⁵ <https://www.data.gouv.fr/fr/>

¹⁶ Directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:FR:PDF>

¹⁷ G8, *Charte du G8 pour l'Ouverture des Données Publiques*, 18 juin 2013. Disponible sur <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/charte-g8-ouverture-donnees-publiques-fr.pdf>

¹⁸ <http://fr.slideshare.net/Etalab/g8-plan-daction-open-data-pour-la-france>

moteur sur l'ouverture des données publiques. C'est par exemple le cas de la question de la gratuité puisque le décret Etalab de 2011 postulait déjà l'absence de redevance par défaut lors de la réutilisation de données publiques.

Citons enfin la circulaire du Premier Ministre du 13 septembre 2013 annonçant la diffusion du *Vademecum sur l'ouverture et le partage des données publiques*¹⁹. Il s'agit d'un guide pédagogique à l'usage des administrations qui explique les enjeux de l'ouverture des données publiques et répond à des questions pratiques : quelles sont les données concernées, quels sont les formats à utiliser, comment documenter les jeux de données ? Il a pour but d'inciter les ministères à référencer leurs données sur la plate-forme *data.gouv.fr*.

Quels sont les bénéfices d'une politique d'*open data* ? La charte du G8 mentionne deux types d'objectifs poursuivis par les politiques d'ouverture des données publiques : améliorer la gouvernance et encourager l'innovation²⁰. Cette vision rejoint celle qui est proposée par Valérie Peugeot, pour qui les arguments en faveur de l'*open data* relèvent à la fois du champ politique ou du champ économique. « L'approche politique insiste sur la revitalisation de la démocratie par la participation, de la citoyenneté par la transparence, et sur l'opportunité de repenser la relation entre administrations et administrés. Tandis que la vision économique évoque l'ouverture des données comme un facteur d'innovation, de création de nouveaux services, de contribution à la croissance, et d'amélioration de la vie quotidienne²¹. »

Afin de conclure sur le contexte des politiques publiques françaises, penchons-nous un instant sur la façon dont se positionnent les tutelles des deux agences bibliographiques par rapport à l'ouverture des données publiques.

Au niveau du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, on trouve des incitations fortes à l'ouverture mais surtout dans le champ des publications (*open access*). Geneviève Fioraso, secrétaire d'État à l'Enseignement supérieur et à la Recherche, affirme ainsi dans son discours du 24 janvier 2013 que « l'information scientifique est un bien commun, qui doit être disponible pour tous²² ». Par ailleurs le monde de la recherche est confronté à l'émergence d'une incitation à ouvrir les données produites dans le cadre de projets de recherche afin de permettre leur diffusion et leur réutilisation. Elle apparaît par exemple dans le cadre d'*Horizon 2020*, le programme européen pour la recherche et le développement pour la période 2014-2020²³.

D'autre part le ministère a ouvert en avril 2014 son propre portail *open data* dédié à des jeux de données concernant l'Enseignement supérieur et la Recherche²⁴. Néanmoins les données des bibliothèques n'y figurent pas et l'ABES n'y est pas référencée.

Au niveau du ministère de la Culture et de la Communication, on entend fréquemment un discours très favorable à l'ouverture des données publiques. Des initiatives œuvrant à ouvrir et rapprocher les référentiels produits par les différentes instances du ministère ont été prises, telles que le programme HADOC (Harmonisation de la production des données culturelles) visant le rapprochement de toutes les

¹⁹ ETALAB, *Vade-mecum sur l'ouverture et le partage des données publiques*, septembre 2013. Disponible sur <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf>

²⁰ Charte du G8, *op.cit.*

²¹ Valérie Peugeot, citée par CHIGNARD Simon, *op.cit.*

²² <http://www.enseignementsup-recherche.gouv.fr/cid66992/discours-de-genevieve-fioraso-lors-des-5e-journees-open-access.html>

²³ <http://www.horizon2020.gouv.fr/>

²⁴ <http://data.enseignementsup-recherche.gouv.fr/explore/>

institutions culturelles par les technologies du web sémantique²⁵. On peut également mentionner la publication de rapports émis par le Département des programmes numériques du ministère²⁶ qui font preuve d'un grand volontarisme en matière d'*open data* pour les données culturelles.

Il faut cependant nuancer cet enthousiasme en mentionnant l'existence de réelles divergences parmi les communautés professionnelles du secteur culturel, notamment liées à la crainte de la perte de redevances, telles qu'elles sont mentionnées dans le rapport Trojette²⁷. D'autre part la relation aux notices descriptives dans les musées ou les services d'archives n'est pas la même que dans les bibliothèques, elles peuvent en effet y être considérées comme des œuvres relevant du droit d'auteur.

La préoccupation d'ouvrir les données culturelles se manifeste également dans le rapport Lescure²⁸, commandé par le ministère de la Culture et de la Communication en 2012. Portant comme titre *Acte II de l'exception culturelle*, ce rapport vise à proposer une adaptation des dispositifs favorisant l'exception culturelle à l'ère du numérique. Dans un chapitre consacré aux métadonnées, le rapport dresse le constat que les bases de métadonnées culturelles sont dispersées, cloisonnées et souvent redondantes. Cela concerne notamment le secteur du livre dans lequel plusieurs acteurs constituent des bases de données bibliographiques, au premier rang desquels la BnF et Électre, auxquels il faut rajouter le Fichier exhaustif du livre alimenté par Dilicom, le Sudoc, catalogue collectif des bibliothèques de l'enseignement supérieur géré par l'ABES, ou encore les opérateurs commerciaux internationaux comme Amazon.

Le rapport indique que cette situation est identique dans les secteurs de la musique ou du cinéma. Il est de fait difficile de disposer de données fiables de référence, ce qui est préjudiciable notamment pour la rémunération des créateurs et ne favorise pas l'émergence d'offres légales. L'une des 80 propositions du rapport Lescure consiste donc à préconiser la création de registres ouverts de métadonnées associant les acteurs du secteur public et du secteur privé²⁹.

Le mouvement de l'*open data* rencontre un autre mouvement, plus ancien, qui est celui-là propre aux bibliothèques. Celles-ci sont en effet engagées dans une transition consistant à faire évoluer leurs dispositifs de signalement afin de les

²⁵<http://www.culturecommunication.gouv.fr/Ressources/Harmonisation-des-donnees-culturelles/Presentation-du-Programme/Presentation-du-programme>

²⁶ Ministère de la Culture et de la Communication, *Ouverture et partage des données publiques culturelles. Pour une (r)évolution numérique dans le secteur culturel*, décembre 2013. Disponible sur <http://www.ladocumentationfrancaise.fr/rapports-publics/144000037/index.shtml>

Ministère de la Culture et de la Communication, *Guide Data Culture : Pour une stratégie numérique de diffusion et de réutilisation des données publiques numériques du secteur culturel*, mars 2013. Disponible sur <http://www.culture.fr/Professionnels/Reutilisation-de-ressources-Culture/Guide-Data-Culture>

²⁷TROJETTE Mohammed Adnène, *Rapport au Premier ministre. Ouverture des données publiques. Les exceptions au principe de gratuité sont-elles toutes légitimes ?* juillet 2013. Disponible sur <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/134000739/0000.pdf>

Remis le 5 novembre 2013, ce rapport analyse des données publiques qui restent soumises à des redevances et souligne l'effet de barrière que constituent ces redevances, mêmes faibles, en particulier pour les réutilisateurs disposant de peu de moyens tels que les citoyens ou les jeunes entreprises. "Cette situation est préjudiciable, au regard des gains attendus d'une politique d'ouverture des données publiques, dont la vocation est la fourniture d'un bien public, vecteur d'externalités positives".

²⁸ LESCURE Pierre, *Mission « Acte II de l'exception culturelle » : contribution aux politiques culturelles à l'ère numérique*, Ministère de la Culture et de la Communication, mai 2013. Disponible sur http://www.culturecommunication.gouv.fr/var/culture/storage/culture_mag/rapport_lescur/index.htm

²⁹ « Il est donc proposé de créer, dans chaque secteur, un registre ouvert de métadonnées, grâce à la coopération de toutes les entités, publiques ou privées, qui détiennent des données pertinentes, et en premier lieu des sociétés de gestion collective. »

rendre plus adaptés à l'évolution des usages et à leur nouvel environnement qu'est le web. La production des métadonnées est au cœur de cette mutation.

Les métadonnées de bibliothèques

Les données publiques dont il est question dans cette étude sont les métadonnées produites par les bibliothèques. Les métadonnées « sont de l'information structurée qui décrit, explique, localise ou facilite autrement l'obtention, l'utilisation ou la gestion d'une ressource d'information. Les métadonnées sont souvent appelées données des données ou information sur l'information³⁰. »

Celles-ci sont de différents types et il peut être utile d'en dresser un bref panorama. On distingue :

- Les métadonnées descriptives, qui servent au signalement des objets. Il s'agit des données décrivant les ressources contenues dans les catalogues. Elles comportent les données bibliographiques, qui décrivent les documents, et les données d'autorité, organisées dans des référentiels, qui permettent d'identifier sans ambiguïté les personnes, les noms de lieux, les œuvres, les sujets. C'est sur les métadonnées descriptives que porte essentiellement ce travail.
- Les métadonnées administratives, concernant par exemple la gestion des droits, la traçabilité, la provenance des objets.
- Les métadonnées techniques, qui documentent le processus de création de l'objet, par exemple le processus de numérisation d'un document, et qui sont notamment utiles à la conservation numérique.

D'autres données sont produites par les bibliothèques et peuvent faire l'objet d'un projet *open data*, telles que les métadonnées de gestion, concernant par exemple la fréquentation d'un établissement ou les horaires d'affluence. Ainsi la bibliothèque des Champs libres à Rennes a rendu disponibles ses données statistiques de fréquentation³¹.

Ouverture juridique

Les données peuvent être considérées comme ouvertes juridiquement lorsqu'une licence clarifie les droits et les obligations dont leur utilisation fait l'objet. Les droits concernent l'utilisation des données, leur modification et leur réutilisation, y compris à des fins commerciales. Les obligations sont celles d'attribution, autrement dit le fait de mentionner la source des données, et celle de partage à l'identique, c'est-à-dire dans les mêmes conditions de licence que les données initiales.

Simon Chignard indique que les deux licences les plus utilisées en France sont la Licence Ouverte de l'État, élaborée par la mission Etalab et la licence ODbL (*Open Database Licence*). La Licence Ouverte n'impose que la mention d'attribution. « D'inspiration libérale, elle [est] jugée plus favorable à la création d'une dynamique économique³² ». Elle est compatible avec des licences à l'étranger (*Open Government Licence* britannique) et les autres standards internationaux tels qu'ODC-BY ou CC-BY. C'est la licence choisie par l'ABES et la BnF pour l'ouverture de leurs données. La licence ODbL quant à elle impose l'obligation de partage à l'identique.

³⁰ « Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information. » NISO, *Understanding Metadata*, 2004. Disponible sur <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

³¹ CARTIER Aurore, *op.cit.*

³² CHIGNARD Simon, *op.cit.*

Lors d'un projet d'ouverture des données, il est toujours préférable de ne pas créer une licence sur mesure mais de choisir une licence parmi celles qui sont disponibles, de manière à ne pas brouiller l'offre et à favoriser l'interopérabilité avec les autres acteurs.

Ouverture technique

Les données peuvent être considérées comme ouvertes sur le plan technique lorsqu'elles sont accessibles de manière structurée, notamment de manière à être exploitables par des machines et dans un format non propriétaire.

Une définition de ce qu'est une donnée publique ouverte a été proposée en 2007 par l'*Open Government Working Group*³³, une association de penseurs du web. Cette définition associe en réalité des critères d'ordre technique, incluant jusqu'aux modalités d'accès à la donnée, mais aussi qualitatif et juridique. Ainsi les données publiques sont considérées comme ouvertes si elles sont mises à disposition selon les huit principes suivants :

- Les données sont complètes : il s'agit de donner accès au jeu de données dans son intégralité.
- Les données sont brutes, primaires.
- Les données sont maintenues à jour.
- Les données sont accessibles à tous, pour tous les usages, aussi bien par une machine que par un humain.
- Le traitement des données peut être automatisé.
- Leur accès est non discriminant, il ne nécessite pas d'inscription préalable.
- Les données sont disponibles dans un format non propriétaire. Elles peuvent l'être dans plusieurs formats si l'on considère que les formats propriétaires sont les plus largement utilisés.
- Les données sont libres d'utilisation, leur usage n'est pas déraisonnablement restreint par la propriété intellectuelle.

Simon Chignard souligne que cette définition de l'open data est précise et exigeante, et que dans la pratique les données sont plus ou moins ouvertes selon le respect de ces critères³⁴. Pourtant ces éléments restent insuffisants pour définir l'ouverture des données sur le plan technique.

Le W3C³⁵ (*World Wide Web Consortium*), organisme de normalisation du web créé en 1994 par Tim Berners-Lee, établit quant à lui une distinction entre les données ouvertes et les données liées. Il définit les données ouvertes comme des données interopérables sur le plan juridique, « utilisables, réutilisables, et redistribuables librement – sujettes tout au plus aux demandes d'attribution et de partage à l'identique³⁶ », comme ce qui a été décrit pour les licences. Quant à ce qu'il nomme l'interopérabilité technique (que nous appelons dans ce mémoire ouverture technique), elle est réalisée par la notion de données liées, des « données publiées selon des principes destinés à faciliter la création de liens³⁷ ». On parle

³³ <http://opengovdata.org/>

³⁴ CHIGNARD Simon, *op.cit.*

³⁵ <http://www.w3c.fr/>

³⁶ GROUPE D'INCUBATION W3C, *Rapport final du groupe d'incubation "Bibliothèques et web de données"*, rapport du 25 octobre 2011. Disponible en français sur : <http://mediatheque.cite-musique.fr/MediaComposite/ARTICLES/W3C/XGR-11d-fr.html>

³⁷ *ibidem*

donc de données liées, *linked data* en anglais, ou du web de données. La notion de *linked open data* désigne l'ensemble des données effectivement liées sur le web et ouvertes juridiquement. Le web sémantique est l'ensemble des standards et technologies développées par le W3C utiles à la mise en œuvre du web de données. L'ouverture des données des bibliothèques passe donc par l'adoption des standards du web sémantique, sur lesquels nous reviendrons en détail.

LES INSTITUTIONS EN PRÉSENCE

DEUX AGENCES BIBLIOGRAPHIQUES

Le paysage institutionnel des bibliothèques françaises est complexe. Sans même évoquer les bibliothèques scolaires, les bibliothèques d'hôpitaux, de prisons, d'associations ou encore de comités d'entreprise, retenons que les bibliothèques de lecture publique relèvent des collectivités territoriales dont elles dépendent (villes, départements et de plus en plus les intercommunalités) mais sont soumises au contrôle technique de l'État, qui s'exerce par l'intermédiaire de l'Inspection générale des bibliothèques, rattachée au Service du livre et de la lecture du ministère de la Culture et de la Communication. Les bibliothèques de l'enseignement supérieur (services communs de la documentation des universités, bibliothèques de grands établissements) relèvent quant à elles du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche. Ces deux catégories de bibliothèques disposent chacune d'une agence bibliographique, même si ce terme recouvre dans les deux cas des réalités différentes. L'agence bibliographique nationale est un département de la BnF, le département de l'information bibliographique et numérique (IBN). La BnF est un établissement public à caractère administratif placé sous la tutelle du ministère de la Culture et de la Communication. L'ABES, également établissement public à caractère administratif, est placée sous la tutelle du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.

La séparation des bibliothèques publiques et des bibliothèques universitaires remonte à 1975, date à partir de laquelle elles ne sont plus administrées par la même tutelle. Rappelons brièvement les étapes de ce processus pour comprendre les origines d'une situation à laquelle il est parfois fait référence sous le nom de « Yalta » entre culture et enseignement supérieur.

A partir de 1945³⁸, date de naissance de la Direction des bibliothèques et de la lecture publique (DBLP), la compétence concernant les bibliothèques et la lecture publique relève du ministère de l'Éducation nationale, alors également en charge des arts et des lettres. En 1959 le ministère des Affaires culturelles est créé et hérite de la Direction générale des arts et des lettres (DGAL). Cependant la DBLP reste dans le giron du ministère de l'Éducation nationale.

En 1975 le décret 75-1003³⁹ du 29 octobre élargit les compétences du ministère de la Culture au détriment du ministère des Affaires étrangères, dessaisi de l'exportation du livre, du ministère de l'Industrie, qui doit renoncer à l'industrie du livre, et du ministère de l'Éducation nationale. Celui-ci perd ses attributions en matière de lecture publique mais seulement en partie, les bibliothèques universitaires et la Bibliothèque nationale restant dans son périmètre. La Bibliothèque nationale passera sous tutelle du ministère de la Culture le 5 juin

³⁸ Comité d'histoire du ministère de la Culture, *Histoire administrative du ministère de la culture et de la communication, 1959-2012. Les services de l'administration centrale*, Paris, La Documentation française, 2012.

³⁹ Décret n°75-1003 du 29 octobre 1975.
http://www.legifrance.gouv.fr/affichTexte.do?jsessionid=528DCE0B6AA8365180D39253753E187A.tpdjo05v_1?cidTexte=JORFTEXT000000500152&categorieLien=id

1981⁴⁰. L'ABES est créée par décret le 24 octobre 1994⁴¹. Les bibliothèques publiques et la BnF d'une part, les bibliothèques de l'enseignement supérieur et l'ABES d'autre part, sont donc depuis sous la responsabilité de deux ministères différents.

On peut légitimement s'interroger sur la pertinence de ce système à deux têtes et se demander s'il ne serait pas plus efficace de ne disposer que d'une seule agence bibliographique pour l'ensemble des bibliothèques du territoire français. Mais « l'idée, un moment caressée par quelques-uns, d'une agence bibliographique nationale, semble aujourd'hui avoir été abandonnée, » indique un rapport de l'IGB (Inspection générale des bibliothèques) de 2011⁴², invitant par ailleurs à amplifier la coopération entre les deux établissements et leurs tutelles. Le rapport n'indique pas pourquoi l'idée d'une agence unique n'est plus d'actualité. Laissons de côté l'aspect politique de cette question et hasardons l'hypothèse que les deux établissements ont peut-être développé une expertise telle dans leurs domaines d'actions respectifs que leur existence à tous deux peut paraître aujourd'hui légitime. Et en effet, ces deux institutions, tout en étant confrontées à des problématiques relativement proches, se distinguent très fortement par leur histoire, leurs missions, leur fonctionnement et leur culture.

LA BIBLIOTHÈQUE NATIONALE DE FRANCE

Nous ne reviendrons pas sur la longue histoire de la Bibliothèque Nationale, héritière de la bibliothèque royale de Charles V. Évoquons brièvement la naissance du nouvel établissement créé en 1994.

En 1988 François Mitterrand annonce son intention d'entreprendre « la construction et l'aménagement de l'une des ou de la plus grande et de la plus moderne bibliothèque du monde.⁴³ » Son conseiller Jacques Attali avait en tête une bibliothèque entièrement numérique qui permettrait de donner accès via un écran d'ordinateur à toutes les bibliothèques du monde. Finalement le projet s'éloigne de cette vision, la nouvelle bibliothèque n'ayant rien d'immatériel et contenant à son ouverture déjà 11 millions de documents. Mais la dimension numérique n'est pas oubliée comme en témoigne le projet de bibliothèque numérique Gallica qui voit le jour dès 1997. Depuis, cette dimension s'est encore étoffée avec le développement de la numérisation de masse, l'extension du dépôt légal à l'archivage du web ou encore la préservation numérique.

La BnF naît donc en 1994 de la fusion de la Bibliothèque Nationale et de l'Établissement public de la Bibliothèque de France alors en charge de la construction du nouveau bâtiment sur le site de Tolbiac. Celui-ci, baptisé Bibliothèque François-Mitterrand, a en charge l'ensemble des imprimés, des périodiques et de l'audiovisuel. L'ancien site de la rue de Richelieu abrite désormais les départements spécialisés dédiés aux manuscrits, aux cartes et plans, aux estampes et photographies, aux monnaies et médailles et enfin à la musique. Un chantier de rénovation et de mise aux normes de ce site est en cours depuis 2010. La BnF est aussi constituée de la Bibliothèque de

⁴⁰ Décret n°81-646 du 5 juin 1981 relatif aux attributions du ministre de la culture. http://www.legifrance.gouv.fr/jopdf/common/fo_pdf.jsp?numJO=0&dateJO=19810606&numTexte=&pageDebut=01700&pageFin=

⁴¹ Décret no 94-921 du 24 octobre 1994 portant création de l'Agence bibliographique de l'enseignement supérieur. http://www.legifrance.gouv.fr/affichTexte.do?jsessionid=D106C38CA846A5EB781E67ADC795878C.tpdjo08v_1?cidTexte=JORFTEXT000000550131&categorieLien=id

⁴² GIRARD Christine, LECOQ Benoît, L'avenir des catalogues collectifs, Inspection Générale des Bibliothèques, rapport n°2011-017, juillet 2011. Disponible sur : http://cache.media.enseignementsup-recherche.gouv.fr/file/2011/86/1/Rapport_cat_coll.-V-25-11-2011_200861.pdf

⁴³ François Mitterrand, TF1, 14 juillet 1988. Retranscription de l'interview par Yves Mourousi disponible sur <http://discours.vie-publique.fr/notices/887019900.html>

l'Arsenal, de la Bibliothèque du musée de l'Opéra et de la Maison Jean Vilar à Avignon, ainsi que des sites techniques de Sablé-sur-Sarthe et Bussy-Saint-Georges.

La BnF est une institution ancienne qui tout en devant veiller à remplir ses missions traditionnelles et à servir son réseau, est confrontée à de nouveaux défis.

Missions

Les missions de la BnF sont stipulées dans son décret de création n°94-3 du 3 janvier 1994. Les deux principales sont :

- « Collecter, cataloguer, conserver et enrichir dans tous les champs de la connaissance, le patrimoine national dont elle a la garde ».
- « Assurer l'accès du plus grand nombre aux collections ».

A ce titre elle procède à la collecte de la production éditoriale française par le biais du dépôt légal et réalise des acquisitions dans la production étrangère. Elle publie la bibliographie nationale française, qui décrit les ouvrages entrés par le biais du dépôt légal. Elle communique ses enregistrements bibliographiques pour faciliter le travail des autres établissements et collabore avec eux pour l'établissement de catalogues collectifs.

La mise à disposition des collections passe par la gestion d'un certain nombre de catalogues, parmi lesquels le catalogue général de la BnF⁴⁴, le catalogue des archives et manuscrits⁴⁵, Mandragore⁴⁶ (base iconographique du département des Manuscrits), le catalogue des médailles et antiques⁴⁷, le catalogue de La Joie par les livres⁴⁸ (catalogue des collections conservées par le Centre national de la littérature pour la jeunesse), etc.

La conservation du patrimoine national pour les générations à venir implique la prise en compte les difficultés dues à la conservation physique des supports. Cette préoccupation concerne tout autant les supports papier (elle se traduit par l'activité des centres de restauration, de désacidification et de microfilmage de la BnF) que le numérique, comme en témoigne le projet SPAR (Système de préservation et d'archivage réparti) lancé en 2008.

L'accès aux collections implique de mettre les ouvrages à la disposition du public, de proposer des expositions, mais aussi de mettre en œuvre des programmes de numérisation, de recherche et de coopération. C'est dans ce sens que la BnF dispose d'un réseau. Avant d'évoquer ces programmes, rappelons qu'il n'existe pas de catalogage partagé parmi les bibliothèques publiques, pratique que la BnF aurait dans le cas contraire toute légitimité à animer.

⁴⁴ <http://catalogue.bnf.fr/>

⁴⁵ <http://archivesetmanuscrits.bnf.fr/>

⁴⁶ <http://mandragore.bnf.fr>

⁴⁷ <http://medaillesetantiques.bnf.fr/>

⁴⁸ <http://lajoieparleslivres.bnf.fr/>

Le réseau de la BnF

L'échec du catalogage partagé

Il faut pourtant souligner qu'il y a eu dans les années 80 une tentative de travail en réseau des bibliothèques publiques. La Direction du livre et de la lecture avait en effet lancé en 1982 un projet nommé LIBRA (logiciel intégré pour les bibliothèques en réseau automatisé)⁴⁹. Il s'agissait également d'un système de catalogage partagé, qui fédérait en 1986 une cinquantaine de bibliothèques, majoritairement des bibliothèques centrales de prêt (qui deviennent alors bibliothèques départementales de prêt) et quelques bibliothèques municipales. Le Centre national de coopération des bibliothèques publiques (CNCBP) est créé en 1986 à Massy pour administrer ce réseau et proposer une offre de formation.

Mais ce réseau est un échec et disparaît en 1989 tandis que le CNCBP est dissous en 1991. Cette décision du ministère s'appuie sur les conclusions d'un schéma directeur de l'information bibliographique, rendues en 1989, selon lesquelles il revenait à la Bibliothèque nationale de produire les notices des ouvrages publiés en France, conformément à sa mission de dépôt légal. Elle est donc considérée comme « le producteur institutionnel de l'information bibliographique française. Cette mise au clair concernant les acteurs en matière de production a conduit la Direction du livre et de la lecture à mettre fin au réseau de catalogage partagé, LIBRA.⁵⁰ »

Cet échec d'une tentative de catalogage en réseau s'explique sans doute en partie par le mouvement de décentralisation initié à partir de 1982. Les bibliothèques font désormais leurs propres choix d'équipement informatique, et se dotent dès 1986 de nouveaux outils de catalogage, notamment pour pallier l'insuffisance de LIBRA à gérer correctement le prêt⁵¹. Il devient dès lors difficile à l'administration centrale de trouver des leviers pour imposer un quelconque mode de fonctionnement. La situation sera différente pour les établissements de l'enseignement supérieur et de la recherche, qui restent dépendants de leur ministère de tutelle, du moins jusqu'à la loi LRU en 2007⁵².

Pôles associés et projets de coopération documentaire

La notion de réseau à la BnF s'entend donc dans le cadre de sa politique de coopération, inscrite dans ses missions. Elle se traduit dans le réseau des pôles associés mais aussi de plus en plus dans une logique de projets autour de programmes de signalement ou de numérisation.

Le réseau des pôles associés de dépôt légal rassemble les bibliothèques de dépôt légal imprimeur (BDLI). Les imprimeurs sont en effet soumis à l'obligation de déposer un exemplaire de chaque document imprimé dans la BDLI de leur région. Ce dispositif qui vient compléter le dépôt légal éditeur, permet de collecter une documentation régionale qui échappe parfois au circuit éditorial traditionnel. Les BDLI sont également des partenaires de la BnF pour la collecte du web. Il s'agit notamment de collecter les sites ayant trait aux élections locales ou à des thématiques patrimoniales régionales. A la

⁴⁹ DUCHEMIN Pierre-Yves, avec la participation de Dominique LAHARY, L'art d'informatiser une bibliothèque. Guide pratique, Paris, Ed. du Cercle de la Librairie, 2000

⁵⁰ LEMELLE Françoise. « Le serveur bibliographique national », in *Bulletin des bibliothèques de France*, n° 3, mai 1991. Disponible sur <http://bbf.enssib.fr/consulter/bbf-1991-03-0224-002>

⁵¹ DUCHEMIN Pierre-Yves, *op.cit.*

⁵² Loi n° 2007-1199 du 10 août 2007 relative aux libertés et responsabilités des universités. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000824315>

suite d'un arrêté du 16 septembre 2014⁵³, la consultation des archives du web jusque-là réservée à la BnF est autorisée dans les 26 BDLI françaises.

La BnF anime également le réseau des pôles associés documentaires, thématiques ou régionaux, composé d'établissements qui détiennent des collections présentant un intérêt national. Elle mène des projets de coopération documentaire concernant le signalement de fonds patrimoniaux, notamment dans le cadre du Plan d'action pour le patrimoine écrit (PAPE), et la numérisation de fonds imprimés. Ces projets couvrent l'ensemble du territoire français et les images et les métadonnées produites alimentent la bibliothèque numérique Gallica.

Citons enfin le partenariat de la BnF avec la Bibliothèque nationale universitaire de Strasbourg autour de Numistral⁵⁴, dans le cadre de l'offre Gallica Marque blanche. Dans ce projet, la BnF met à disposition de la BNU une infrastructure de bibliothèque numérique reposant sur Gallica, personnalisée selon une charte graphique propre à la BNU. Les documents de Numistral enrichissent en retour les collections de Gallica.

L'ensemble du réseau de la BnF réunit près de 200 établissements, de tous types et de tous statuts : bibliothèques de lecture publique mais aussi de l'enseignement supérieur et de la recherche, de ministères ou d'associations.

La mise à disposition de produits bibliographiques

La BnF peut également être considérée comme une tête de réseau par la diffusion de ses produits bibliographiques tels que la bibliographie nationale, mais aussi par la diffusion directe de ses notices bibliographiques et d'autorité. Celles-ci sont en effet mises à la disposition des bibliothèques qui souhaitent les utiliser pour leur catalogue. Cette pratique, appelée la dérivation de notices, permet d'alléger considérablement les tâches de catalogage. Il est possible d'obtenir rétrospectivement un ensemble de notices (depuis 1970) ou uniquement les notices courantes (du mois précédent). Ce service est gratuit, seuls les produits nécessitant une intervention personnalisée sont facturés.

Malheureusement ce service n'est pas utilisé autant que l'on pourrait l'espérer. En 2010 seulement 35 % des bibliothèques municipales pratiquent la dérivation de notices⁵⁵. Ceci est dû en partie à des difficultés techniques ou à un manque de formation ou d'information. Mais le principal obstacle est que les bibliothèques ont besoin de disposer des notices dès la parution des documents, au moment où elles en font l'acquisition. Or les notices de la BnF sont souvent disponibles trop tard, le délai engendré par le processus du dépôt légal étant de plusieurs semaines. La BnF a mené une enquête en 2012 sur les pratiques des bibliothèques⁵⁶ afin de mettre en place des actions pour améliorer le taux de récupération de ses notices.

La BnF anime donc un réseau de pôles associés, elle diffuse ses produits bibliographiques, mais on ne peut pas dire qu'elle ait un réseau au même titre que celui qu'anime l'ABES. Le catalogue général de la BnF reflète uniquement les collections de l'établissement. On pourrait objecter l'existence du Catalogue

⁵³[http://legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?](http://legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=20141007&numTexte=27&pageDebut=16263&pageFin=16263)

[numJO=0&dateJO=20141007&numTexte=27&pageDebut=16263&pageFin=16263](http://legifrance.gouv.fr/jopdf/common/jo_pdf.jsp?numJO=0&dateJO=20141007&numTexte=27&pageDebut=16263&pageFin=16263)

⁵⁴ <http://www.numistral.fr/>

⁵⁵ Ministère de la Culture et de la Communication, Boîte à outils du numérique en bibliothèque, Fiche pratique n°1, La récupération des données pour le catalogue, 2012. Disponible sur <http://www.culturecommunication.gouv.fr/Disciplines-secteurs/Livre-et-Lecture/Bibliotheques/Numerique-et-bibliotheques/Boite-a-outils-du-numerique-en-bibliotheque>

⁵⁶ http://www.bnf.fr/documents/recuperation_donnees_bnf_enquete2013.pdf

Collectif de France (CCFr), mais celui-ci ne fait qu'agréger des catalogues existants, il n'est pas un outil de production partagée. La BnF possède une forte culture de la diffusion mais elle laisse difficilement entrer ce qui est produit à l'extérieur. Penchons-nous maintenant sur ses fonctions d'agence bibliographique nationale.

L'agence bibliographique nationale

Le contrôle bibliographique universel

La notion d'agence bibliographique nationale doit être comprise dans un contexte international qui est celui du contrôle bibliographique universel (CBU). Dorothy Anderson en a formulé les principaux principes dans un article paru en 1974⁵⁷, décrivant un système né en réalité au début des années soixante. Le CBU repose sur l'idée que chaque pays est le mieux placé pour identifier et signaler ses propres publications. En appliquant les normes bibliographiques internationales, il est possible grâce à l'action de chaque pays de disposer du signalement de l'ensemble de la production éditoriale mondiale : « *The purpose of the system is to make universally and promptly available, in a form which is internationally acceptable, basic bibliographic data on all publications issued in all countries*⁵⁸. »

Ce système repose entièrement sur la publication, à une époque où l'accès à l'information passe obligatoirement par un document imprimé. Dans la pratique, ce système se traduit par un échange de notices bibliographiques établies sur fiches, retranscrites par des dactylographes et envoyées par paquets aux quatre coins du monde. Au moment de l'informatisation des catalogues, ce procédé sera remplacé par l'échange de notices au format MARC.

Pour être opérationnel, le contrôle bibliographique international comporte des outils à déployer à l'échelle nationale. Ces outils sont l'agence bibliographique nationale, généralement établie au sein de la bibliothèque nationale, le dépôt légal, la bibliographie nationale, qui recense et décrit les publications entrées dans le cadre du dépôt légal et qui est publiée périodiquement. On peut ajouter à ces outils la gestion des fichiers d'autorité nationaux mais aussi les fonctions de normalisation.

Du CBN au département de l'IBN

C'est en 1976 qu'une unité regroupe pour la première fois à la BnF toutes les fonctions d'agence bibliographique⁵⁹, avec la création du Centre bibliographique national (CBN), deux ans après la publication du manifeste de Dorothy Anderson. Ce moment coïncide également avec les débuts de l'informatisation de la bibliographie nationale.

Cette unité devient en 1985 le Centre de coordination bibliographique et technique (CCBT) qui prend désormais en compte la gestion des fichiers d'autorités de la Bibliothèque nationale, qui viennent d'apparaître en respect des règles énoncées en 1974. En 1994, avec la naissance de la BnF, une réorganisation donne lieu à un nouvel organigramme contenant désormais une unité nommée Agence bibliographique nationale (ABN). En 2008 a lieu la fusion de l'Agence bibliographique nationale et du pôle Information numérique du département de la Bibliothèque numérique, chargé de la production et de la diffusion des documents numériques ainsi que de la collecte du web.

⁵⁷ ANDERSON Dorothy, *Universal Bibliographic Control. A long term policy. A plan for action*, PullachbeiMünchen, VerlagDokumentation, 1974

⁵⁸ *op.cit.*

⁵⁹ Cela ne signifie pas que ces fonctions sont nouvelles, comme en atteste l'existence du dépôt légal depuis 1587.

Cette fusion aboutit à la formation d'une nouvelle entité, le département de l'Information bibliographique et numérique (IBN). Le numérique a fait officiellement son entrée dans l'identité de l'agence bibliographique, traduisant un glissement de la gestion des catalogues vers celle des métadonnées.

Le département de l'IBN⁶⁰ continue à jouer le rôle d'agence bibliographique, en assurant la coordination des missions suivantes :

- Le dépôt légal, réalisé dans d'autres départements de la BnF⁶¹. Son directeur préside le Comité de coordination du dépôt légal, qui a en particulier pour missions le suivi de l'évolution de la législation.
- La diffusion des données, dont la publication de la *Bibliographie nationale française*.
- La normalisation.
- La modélisation des données.
- Les fichiers d'autorité (personnes, collectivités, titres).
- Le centre national Rameau, responsable de la gestion intellectuelle du langage d'indexation Rameau (Répertoire d'autorité-matière encyclopédique et alphabétique unifié).
- Le centre national ISSN, chargé de l'attribution d'identifiants ISSN (*International Standard Serial Number*) aux publications en série françaises.
- La communication et la formation professionnelle.
- Le contrôle qualité et la coordination de la production dans les autres départements.

Actualité de la BnF

La BnF a formalisé sa stratégie pour la période 2014-2016 dans un contrat de performance⁶². Ce document comporte de nombreux axes que nous n'analyserons pas dans leur totalité. Certains objectifs qui y sont présentés sont naturellement dans la continuité des missions traditionnelles de la BnF. Ils concernent les acquisitions, courantes ou patrimoniales, l'augmentation de la fréquentation de l'établissement ou encore la rénovation du site Richelieu. D'autres objectifs liés à la maîtrise des dépenses nous rappellent que la BnF évolue comme tous les services publics dans un contexte budgétaire contraint. Mais nous souhaiterions ici extraire quelques points de ce contrat qui nous paraissent illustrer tout particulièrement l'actualité à laquelle est confrontée la BnF et qui entrent en résonance avec notre sujet. Ces points pourraient se résumer par la volonté de la BnF de s'adapter à la « transition numérique », ainsi que cela est mentionné dans

⁶⁰ Dans le cadre de cette étude, qui consiste à analyser la coopération entre l'ABES et la BnF, il sera souvent question, plutôt que la BnF dans son ensemble, du département de l'IBN qui est aujourd'hui l'entité principale par laquelle s'effectue la collaboration avec l'ABES.

⁶¹ « Le département du Dépôt légal pour les documents imprimés [...] et les documents numériques en ligne, le département de l'Audiovisuel pour les documents sonores, audiovisuels, multimédias et les documents numériques sur support, le département des Estampes et de la photographie [...], le département de la Musique [...], le département des Cartes et plans ».

[http://www.bnf.fr/fr/professionnels/depote_legal_definition/s.depote_legal_organisation.html?first_Art=non](http://www.bnf.fr/fr/professionnels/depot_legal_definition/s.depote_legal_organisation.html?first_Art=non)

⁶² BnF, *Contrat de performance 2014-2016*, mai 2014. Disponible sur http://www.bnf.fr/fr/la_bnf/connaitre_bnf/a.contrat_performance.html

l'introduction du document. Ce concept nous rappelle celui de transition bibliographique qui sera développé un peu plus loin et qui est lui-même fortement lié à l'émergence du numérique.

Faire évoluer le dépôt légal

L'une des grandes actualités de la BnF concerne l'adaptation de la mission de dépôt légal au numérique. Ce mouvement a été initié avec le dépôt légal du web, pour lequel des expérimentations ont été entreprises dès 1999. Il a été conforté en 2006 à la suite de l'adoption de la loi DADVSI⁶³ qui a mis en place le dépôt légal du web et de son décret d'application du 19 décembre 2011⁶⁴.

Il s'agit désormais d'adapter le dépôt légal aux livres numériques, ce qui nécessite la mise en place d'une chaîne de traitement appropriée mais aussi un ajustement du cadre réglementaire. L'objectif que s'est fixé la BnF consiste à avoir étendu le champ de dépôt légal à l'ensemble de l'édition numérique en 2016.

Il faut également mentionner une évolution du dépôt légal traditionnel qui va dans le sens d'une simplification de la procédure. En 2009 a été mis en place l'extranet du dépôt légal, qui permet aux éditeurs de saisir en ligne leurs déclarations au lieu d'envoyer des formulaires que la BnF devait ressaisir. Dans la continuité de cette fluidification de la chaîne, un projet est en cours consistant à proposer aux éditeurs déposant le plus de publications de transférer directement leurs données à la BnF depuis leur système d'information.

Il s'agit donc de mettre à profit le dépôt légal pour obtenir les flux de données bibliographiques directement auprès des éditeurs, ce qui représente une économie de ressources du côté de la BnF, mais aussi une piste de raccourcissement des délais de création des notices. On l'a vu plus haut, ce point pourrait avoir un impact positif sur la dérivation de notices par les bibliothèques réutilisatrices, favorisant là encore une économie de catalogage. On voit ainsi se dessiner une tendance qui consiste à favoriser autant que possible la réutilisation des métadonnées au plus près de leur source de production. D'autre part ce système permettra d'obtenir via ces flux de données au format ONIX, convertis en MARC, des données destinées à l'origine aux librairies, telles que les visuels de couverture, les résumés, des photos ou biographies des auteurs, qui viendront enrichir le catalogue de la BnF et le rendre plus attractif.

Dérivation des notices d'OCLC

Dans la même perspective la BnF récupère elle aussi des notices bibliographiques auprès d'OCLC (*Online Computer Library Center*). OCLC est une organisation fédérant des bibliothèques du monde entier, basée dans l'Ohio, qui a un positionnement particulier puisque tout en étant une association à but non lucratif au service des bibliothèques, OCLC est devenu « un acteur incontournable avec la posture d'une multinationale⁶⁵ ». Tête de réseau de catalogage, fournisseur de bases de données, éditeur de logiciels, développeur de services web à destination des bibliothèques, OCLC gère également le catalogue mondial Worldcat, qui contient 312 millions de notices issues des 121 bibliothèques membres. La BnF en fait partie depuis 2009 et en retour utilise les notices de Worldcat pour ses acquisitions courantes en langue étrangère.

⁶³ LOI n° 2006-961 du 1er août 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000266350&dateTexte=&categorieLien=id>

⁶⁴ Décret n° 2011-1904 du 19 décembre 2011 relatif au dépôt légal. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000025002022&dateTexte=&categorieLien=id>

⁶⁵ BOURDENET Philippe, « OCLC, l'histoire d'une coopération fructueuse », in *Documentaliste-Sciences de l'information*, vol.50, n°2, juillet 2013, p.30

Cette pratique de dérivation des notices du réservoir OCLC doit encore être consolidée selon le contrat de performance de la BnF qui fixe l'objectif de passer à 95 % de notices dérivées en langue étrangère en 2016. Cet indicateur s'accompagne d'un objectif de réduction du délai de catalogage. « Il s'agit ici de favoriser les économies et la mutualisation dans les activités de catalogage, ainsi que l'interopérabilité avec les autres producteurs ou détenteurs de données bibliographiques »⁶⁶.

Une nouvelle interface pour le catalogue

Un autre point évoqué dans le contrat de performance de la BnF est celui d'une nouvelle interface publique pour le catalogue. L'objectif est d'offrir une navigation plus intuitive, un tri des résultats par facettes, des fonctions permettant de rebondir vers des ressources complémentaires. Il s'agit de prendre en compte les habitudes de navigation des utilisateurs prises sur le web en permettant d'effectuer des découvertes, en favorisant une plus grande sérendipité. Une première version de cette interface a été mise en ligne en octobre 2014, permettant aux agents de la BnF mais aussi au public de faire remonter des commentaires à l'équipe de développement⁶⁷. L'objectif est d'adopter définitivement cette nouvelle version de l'interface du catalogue général en 2016.

Le projet ReLire

Le projet ReLire s'appuie sur l'idée que le numérique peut redonner une vie aux œuvres qui ne sont plus publiées, tout en prenant en compte la difficulté propre aux ouvrages qui sont encore soumis au droit d'auteur, celui-ci s'appliquant 70 ans après la mort de l'auteur. Pour répondre à cela la loi du 1^{er} mars 2012⁶⁸ a préconisé la création du Registre des livres indisponibles en réédition électronique (ReLire). La publication de la liste des livres qui sont encore sous droit d'auteur et qui ne sont plus disponibles dans le circuit commercial, sous forme imprimée ou numérique, a été confiée à la BnF. Cette liste est mise à jour une fois par an. Dans une seconde phase la BnF procédera à la numérisation des livres qui seront disponibles à la consultation dans Gallica intra muros⁶⁹.

Vers la transition bibliographique

Enfin le contrat de performance mentionne de nombreux points confortant la BnF dans son rôle d'agence bibliographique nationale et la positionnant comme un acteur majeur dans la transition bibliographique : mise en œuvre des nouvelles normes de catalogage (RDA), consolidation et enrichissement des référentiels d'autorité, évolution des identifiants. Ces points, qui seront développés dans la suite de cette étude, sont constitutifs d'une stratégie de développement axée plus largement sur l'ouverture et la dissémination des données de la BnF.

L'un des objectifs indiqués dans ce cadre concerne en outre la coopération avec l'ABES, via « l'élaboration d'un plan d'action intégrant l'évolution et la complémentarité des missions des deux opérateurs⁷⁰. » Le moment est donc venu de nous pencher sur la seconde agence bibliographique française.

⁶⁶ BnF, *Contrat de performance 2014-2016*, op.cit.

⁶⁷ <http://cataloguelabs.bnf.fr/index.do>

⁶⁸ LOI n° 2012-287 du 1er mars 2012 relative à l'exploitation numérique des livres indisponibles du XXe siècle. <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000025422700&categorieLien=id>

⁶⁹ Gallica intra-muros est constituée de la partie publique de Gallica à laquelle s'ajoutent les documents accessibles uniquement dans les murs de la BnF (documents sous droits d'auteur, archives du web).

⁷⁰ BnF, *Contrat de performance 2014-2016*, op.cit.

L'AGENCE BIBLIOGRAPHIQUE DE L'ENSEIGNEMENT SUPÉRIEUR

La naissance de l'ABES

Création de l'agence

Contrairement à l'agence bibliographique nationale qui est héritière de la longue histoire de la BnF, l'ABES est un établissement récent. Elle est née de la volonté de créer le catalogue collectif Sudoc (Système universitaire de documentation), dans un contexte où coexistaient dans le champ des bibliothèques de l'enseignement supérieur une hétérogénéité de réseaux et d'applications :

- Les trois réseaux informatisés de catalogage BN-Opale, Auroc (partie française du réseau OCLC) et Sibil. Les données produites par ces réseaux alimentaient le Pancatalogue, catalogue collectif des ouvrages.
- Le catalogue collectif national (CCN) pour les périodiques.
- La base Téléthèses pour les thèses soutenues en France.

La création de l'ABES s'inscrit dans le cadre d'un schéma directeur informatique des bibliothèques universitaires élaboré en 1992 qui visait à intégrer l'ensemble de ces applications, dans le but de fournir aux étudiants et chercheurs un point d'entrée unique aux ressources documentaires mais aussi de rationaliser les moyens consacrés aux tâches de catalogage.

Le décret n°94-921 du 24 octobre 1994 portant création de l'Agence bibliographique de l'enseignement supérieur précise ses missions⁷¹ :

- Recenser et localiser les fonds documentaires des bibliothèques de l'enseignement supérieur.
- Assurer la coordination du traitement documentaire des collections.
- Assurer la gestion et le développement des systèmes et des applications informatiques.
- Éditer les produits dérivés des catalogues.
- Apporter son concours aux établissements dans le domaine de l'information bibliographique.
- Coopérer avec les organismes concourant aux mêmes fins, en France et à l'étranger.

L'ensemble des applications a donc été regroupé à Montpellier, initialement sous la responsabilité d'une vingtaine de personnes (composée de deux tiers d'informaticiens et d'un tiers de bibliothécaires). Les forces de l'ABES à ses débuts ont été exclusivement tournées vers la création du Sudoc.

Le Sudoc

Le Sudoc agrège l'ensemble des données des applications préexistantes. S'initient également à cette occasion les échanges de données avec la BnF qui fournit à l'ABES

⁷¹ Décret no 94-921 du 24 octobre 1994 portant création de l'Agence bibliographique de l'enseignement supérieur. http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=D106C38CA846A5EB781E67ADC795878C.tpdjo08v_1?cidTexte=JORFTEXT00000550131&categorieLien=id

pour le démarrage du Sudoc l'ensemble de ses données d'autorité. Le cœur du système repose sur CBS, un logiciel du prestataire PICA racheté depuis par OCLC.

Un premier déploiement sur 7 sites pilotes est entrepris en 1998. En 2000 le Sudoc est officiellement lancé et disponible en consultation sur le web. Puis la première phase de déploiement est achevée en 2002 avec 113 établissements qui forment le réseau Sudoc. En effet le Sudoc est à la fois un catalogue collectif et un réseau, en réalité deux réseaux :

- Le réseau Sudoc des bibliothèques déployées, qui cataloguent tous leurs documents dans le Sudoc, comportant aujourd'hui 160 établissements regroupant 1450 bibliothèques et centres de documentation⁷². Chaque année des vagues de déploiement sont réalisées pour intégrer de nouveaux établissements et de nouvelles bibliothèques à l'intérieur du réseau.

- Le réseau Sudoc-PS (pour publications en série), plus vaste, comportant des bibliothèques qui ne cataloguent que leurs périodiques dans le Sudoc. Ce réseau est coordonné par l'ABES via 32 centres régionaux en direction de 2000 bibliothèques de tous statuts (territoriales, privées, associatives...)

Le Sudoc est un outil de consultation accessible sur le web⁷³ qui offre également un accès personnalisé par bibliothèque (le « Sudoc local »). C'est aussi un outil de prêt entre bibliothèques pour les professionnels ou directement par l'utilisateur final. C'est enfin un outil de production mutualisée. Le catalogage est réalisé dans le Sudoc puis les notices sont réintégrées dans les SIGB (systèmes intégrés de gestion de bibliothèque) locaux. Il est possible de recevoir l'ensemble des mises à jour effectuées par le réseau ou bien seulement celles effectuées par son établissement, que l'on appelle les mises à jour propres.

L'animation du réseau

La mise en œuvre d'un tel système repose sur la mise à disposition et la maintenance d'un système d'information mais aussi sur l'animation du réseau. La possibilité de réaliser un catalogage collectif passe en effet par l'utilisation de règles et d'outils communs : le format d'encodage UNIMARC, le langage d'indexation Rameau, des règles de catalogage et des bonnes pratiques communes. Des informations sur tous ces éléments sont diffusées par l'ABES via des outils tels que le guide méthodologique Sudoc⁷⁴, un portail très riche visant à aider les membres du réseau dans toutes leurs tâches.

D'autre part la communication au sein du réseau s'organise via des listes de diffusion permettant à tous les membres de poser des questions ou de partager des informations. L'ABES publie également des blogs spécialisés par sujets : la formation⁷⁵, les aspects techniques⁷⁶, le projet de SGB mutualisé⁷⁷, le code de catalogage RDA⁷⁸.

Un autre mode d'animation du réseau passe par des personnes relais dans les établissements qui sont les interlocuteurs privilégiés de l'ABES. Il s'agit des

⁷² <http://www.abes.fr/Sudoc/Le-reseau-Sudoc/Les-etablissements-membres>

⁷³ <http://www.sudoc.abes.fr>

⁷⁴ <http://documentation.abes.fr/sudoc/>

⁷⁵ ABES, *Le blog de la formation*, [en ligne]. Montpellier, ABES. Disponible sur <http://laformation.abes.fr/>

⁷⁶ ABES, *Punktokomo ; blog technique de l'ABES*, [en ligne]. Montpellier, ABES. Disponible sur <http://punktokomo.abes.fr/>

⁷⁷ ABES, *SGB mutualisé*, [en ligne]. Montpellier, ABES. Disponible sur <http://sgbm.abes.fr/>

⁷⁸ ABES, *RDA@ABES*, [en ligne]. Montpellier, ABES. Disponible sur <http://rda.abes.fr/>

correspondants pour le catalogage ou les autorités, des coordinateurs locaux ou encore des formateurs relais. Ces correspondants se rencontrent lors de journées d'information et d'échanges qui leur sont dédiées. Ils sont également mobilisés par l'ABES lors du lancement de chantiers qualité visant à apporter des corrections aux données.

L'ABES propose également une offre de formation destinée à l'utilisation des outils qu'elle met à disposition du réseau. Elle coordonne une offre de formation en présentiel permettant de s'initier aux différentes applications ou pour se former à certaines fonctions (coordinateur Sudoc, responsable de centre régional Sudoc-PS). Elle dispense également une offre de formation à distance de perfectionnement via la mise à disposition d'une plate-forme d'autoformation qui propose des tutoriels ou encore par le biais des « J.e-cours », des sessions de formation continue consultables en direct ou en différé.

Enfin elle publie une revue trimestrielle, *Arabesques*⁷⁹, et organise tous les ans depuis le lancement du Sudoc les journées ABES, événement qui permet au personnel de l'agence et aux membres du réseau d'échanger sur leurs pratiques professionnelles, de présenter les projets en cours et d'avoir un questionnaire prospectif sur l'avenir des applications.

Bilan

Le Sudoc est aujourd'hui déployé dans toutes les universités françaises. Il est selon le rapport d'évaluation de l'AERES (Agence d'évaluation de la recherche et de l'enseignement supérieur) « une réussite reconnue par la communauté des bibliothèques⁸⁰ ». Il a en effet réussi à fédérer des réseaux hétérogènes et a bien atteint son but de diminution de la tâche de catalogage⁸¹.

On peut en effet concevoir combien est complexe la création d'un réseau de catalogage partagé, dans un contexte où chaque établissement est singulier, possède son SIGB et sa politique documentaire propres, et alors que la vocation de l'ABES n'est pas de s'immiscer dans la gestion locale. Il s'agit donc d'une réussite, même si l'on peut apporter quelques nuances.

En effet le travail en commun n'est pas exploité au mieux. La réception des mises à jour propres est l'option retenue dans 80 % des cas, et de nombreuses mises à jour dans les systèmes locaux ne sont pas répercutées, parfois jusqu'à 30 % du catalogue local⁸². On pourrait également déplorer l'absence de concertation sur le catalogage, alors qu'il semblerait pertinent de répartir les domaines de compétences en fonction des spécialités des établissements.

Si la mise en place du Sudoc a mobilisé l'ABES à ses débuts, celle-ci a peu à peu diversifié ses outils et ses activités.

La diversification des activités de l'ABES

Avant de nous pencher sur le contenu de cette diversification, signalons l'évolution majeure du contexte institutionnel que constitue la loi relative aux libertés et

⁷⁹ <http://www.abes.fr/Arabesques>

⁸⁰ AERES, Rapport d'évaluation de l'Agence bibliographique de l'enseignement supérieur, juillet 2012. Disponible sur : <http://www.aeres-evaluation.fr/Etablissements/AGENCE-BIBLIOGRAPHIQUE-DE-L-ENSEIGNEMENT-SUPERIEUR-ABES>

⁸¹ L'objectif du Sudoc était de réduire le nombre de création de notices à 20 %, les 80 % restants devant être de simples localisations.

⁸² BERNON Jean, « Focus sur le projet SGBm », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canal2.tv/video.asp?idvideo=12650>

responsabilités des universités (LRU) du 10 août 2007. Selon Raymond Bérard, directeur de l'Agence de 2007 à 2013, cette loi « inscrit désormais l'action de l'ABES dans une double exigence : mettre en œuvre une stratégie nationale définie par le Ministère et répondre aux attentes de mutualisation des établissements en concevant des produits et services qui répondent à leurs attentes⁸³. »

Calames

Dès 2007 l'ABES propose en complément du Sudoc un nouveau catalogue, Calames, le Catalogue en ligne des archives et des manuscrits de l'enseignement supérieur. De la même manière qu'à la BnF avec son catalogue BAM (BnF archives et manuscrits), s'est fait ressentir la nécessité d'un outil adapté à la description archivistique, très hiérarchisée, reposant sur le format EAD (*Encoded Archival Description*). Calames, comme le Sudoc, est un catalogue partagé par les établissements de l'enseignement supérieur.

La stratégie des API

A partir de 2008, un changement de politique informatique met en place la « stratégie des API ». Visant à atténuer la dépendance du Sudoc aux outils d'OCLC, cette stratégie repose sur le développement d'interfaces de programmation (*Application Programming Interface*) permettant la création de nouveaux outils. Ces outils s'appuient sur une base miroir du Sudoc, synchronisée avec la base de production. Citons par exemple PÉRISCOPE, outil à destination du réseau Sudoc-PS pour la gestion des plans de conservation partagée ou encore Colodus, qui permet d'intervenir sur les données d'exemplaires des périodiques sans passer par l'interface de catalogage standard (WinIBW).

Un autre de ces outils est l'application IdRef. Lancée en 2010, elle est constituée de la base des autorités du Sudoc (personnes, collectivités, autorités matières Rameau ...) et devient un service autonome réutilisable par d'autres applications de l'ABES et par des acteurs extérieurs. Elle est consultable sur le web en accès public. Elle reste le service pour la création des autorités du Sudoc, mais aussi de Calames et des applications liées aux thèses, présentées ci-après.

La valorisation des thèses

L'ABES joue un rôle croissant dans la valorisation des thèses. Cela débute en 2006 avec Star, à la suite de l'arrêté du 7 août 2006 relatif aux modalités de dépôt, de signalement, de reproduction, de diffusion et de conservation des thèses⁸⁴. Star est un outil de production de métadonnées servant à décrire les thèses soutenues déposées au format électronique. En 2011 Step, outil de signalement des thèses en cours de préparation, remplace le fichier central des thèses, dont la gestion a été transférée à l'ABES en 2009. En 2011 le portail des thèses voit le jour afin de répondre à une recommandation du rapport Salençon⁸⁵ publié en 2008 et qui déplorait le manque de visibilité des thèses soutenues en France.

⁸³ BÉRARD Raymond, « L'Abes à l'horizon 2016 », in Bulletin des bibliothèques de France, n° 1, janvier 2013. Disponible sur le : <http://bbf.enssib.fr/consulter/bbf-2013-01-0072-016>

⁸⁴ <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000635069>

⁸⁵ <http://www.enseignementsup-recherche.gouv.fr/cid21677/rapport-du-comite-ist-information-scientifique-et-technique.html>

L'acquisition de documentation électronique

En 2001 l'ABES commence à effectuer des groupements de commandes et joue un rôle croissant dans l'acquisition de documentation électronique. Un partenariat avec le consortium Couperin⁸⁶ commence en 2004. Celui-ci est en charge de l'évaluation des besoins des établissements et de la négociation auprès des éditeurs, tandis que l'ABES assure le portage administratif et financier des marchés d'acquisition de ressources : elle passe les appels d'offre, exécute les marchés, règle les éditeurs et refacture les établissements.

De nouvelles missions pour l'ABES

En 2010 le ministère de l'Enseignement supérieur et de la Recherche transfère à l'ABES plusieurs missions nationales :

- Le signalement et la valorisation des thèses (qui se traduit par la création du portail des thèses évoqué plus haut).
- La conversion rétrospective des catalogues.
- L'organisation des réseaux et des projets coopératifs (prêt entre bibliothèques, réseau du Sudoc-PS et signalement de la documentation électronique).
- La normalisation documentaire.

Par ailleurs, en 2010 est créée la Bibliothèque scientifique numérique (BSN), un dispositif mis en place pour rendre cohérentes les structures en charge de l'information scientifique et technique en France et favoriser un rapprochement entre les acteurs, en particulier les universités et les organismes de recherche. Les principaux partenaires de l'ABES au sein de BSN sont le CNRS⁸⁷ (Centre national de la recherche scientifique), avec notamment l'INIST⁸⁸ (Institut de l'information scientifique et technique), le CCSD⁸⁹ (Centre pour la communication scientifique directe), Couperin et la BnF. L'action de l'ABES s'inscrit dans le cadre de la BSN, en particulier sur les trois segments dont elle assure le pilotage ou le copilotage : les acquisitions numériques, le signalement des ressources, la fourniture de documents.

Dans ce cadre l'ABES a été mandatée pour négocier et acquérir des corpus de ressources fermés (archives de revues, *e-books*) sous forme de licences nationales. Elle a aussi en charge leur signalement et propose un accès centralisé sur le site dédié aux licences nationales de l'ABES⁹⁰. Le signalement de ces ressources est également assuré dans le Sudoc pour les corpus acquis, et il est possible de les récupérer dans les SIGB locaux. Ces ressources ont aussi vocation à être accessibles via la future plate-forme ISTEEX qui sera opérationnelle courant 2015⁹¹. L'ABES travaille également à rendre possible le signalement de ces ressources dans les bases de connaissances exploitées par les outils de découverte.

⁸⁶ Consortium unifié des établissements universitaires et de recherche pour l'accès aux publications numériques <http://www.couperin.org/>

⁸⁷ <http://www.cnrs.fr/>

⁸⁸ <http://www.inist.fr/>

⁸⁹ <http://www.ecsd.cnrs.fr/>

⁹⁰ <http://www.licencesnationales.fr/>

⁹¹ ISTEEX (Initiative d'excellence de l'Information Scientifique et Technique) s'inscrit dans le cadre du programme « Investissements d'avenir ». « Le projet ISTEEX a pour principal objectif d'offrir, à l'ensemble de la communauté de l'enseignement supérieur et de la recherche, un accès en ligne aux collections rétrospectives de la littérature scientifique dans toutes les disciplines en engageant une politique nationale d'acquisition massive de documentation : archives de revues, bases de données, corpus de textes... » www.istex.fr

En 2013 la BSN a également donné mandat à Couperin pour l'abonnement à des ressources courantes en licences nationales (*Freedom Collection* d'Elsevier). L'accord qui a été signé pour les 5 ans à venir a suscité de nombreuses réactions⁹².

Un projet d'établissement

Ces nouvelles missions ont donné une autre dimension à l'ABES. D'opérateur technique elle est devenue une agence de pilotage. Cette transition s'est accompagnée d'un nouveau mode de fonctionnement pour l'agence. Une fois le projet du Sudoc accompli, le besoin d'un projet d'établissement pour structurer son action s'est fait sentir. Le premier projet d'établissement de l'ABES voit le jour pour la période 2008-2011.

Le projet d'établissement 2012-2015⁹³ s'inscrit à la fois dans la continuité des missions initiales (« consolider les services et le réseau », « faire évoluer les catalogues et applications ») et dans ses nouvelles missions (« signaler les ressources des universités et établissements », « développer l'achat de ressources numériques »). Il fait mention également de nouveaux projets dont certains sont des axes à part entière du projet d'établissement (« proposer aux établissements un système de gestion mutualisé », être « partenaire du hub des métadonnées de la recherche »).

Les nouveaux projets

Bacon

Le projet Bacon⁹⁴, pour Base de connaissances nationale, est en cours depuis septembre 2013 en partenariat avec le consortium Couperin.

Une base de connaissances est une base de données qui rassemble les données bibliographiques des ressources imprimées ou numériques, accessibles sur place ou à distance, payantes ou non. Les bases de connaissances sont créées par des sociétés commerciales et sont l'un des éléments constitutifs des outils de découverte, ceux-ci étant «des outils destinés aux bibliothèques, conçus pour fournir un seul point d'accès et indifféremment du support, du mode et du lieu de stockage, à toutes les ressources qu'elles proposent »⁹⁵.

L'étude menée par l'ABES et Couperin a permis de constater que les différents outils disponibles sur le marché fournissaient des données incomplètes et incohérentes. Pour remédier à ce problème les établissements cherchent à obtenir les bonnes informations auprès des éditeurs et des fournisseurs de bases de connaissances, ce qui occasionne pour eux des efforts importants et difficiles à mutualiser. L'étude a donc conclu à la nécessité de fournir aux établissements un « service de données de référence », fiable, consacré en priorité aux bouquets

⁹² Pour de nombreuses raisons, notamment son coût élevé dans un contexte de restriction budgétaire ou encore certaines clauses telles que celle concernant le *data mining*, par laquelle Elsevier revendiquait en quelque sorte un droit d'auteur sur les données brutes. Pour un aperçu des débats sur ce sujet, voir http://corist-shs.cnrs.fr/AccordElsevier_questions

⁹³ ABES, *Projet d'établissement 2012-2015*, novembre 2011. Disponible sur <http://www.abes.fr/Connaitre-l-ABES/Projet-d-etablissement>

⁹⁴ Nous nous appuyons pour décrire ces différents projets sur les présentations données aux journées ABES 2014, en l'occurrence pour Bacon : BOBER Benjamin, « Focus sur BACON », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canalc2.tv/video.asp?idvideo=12652>

⁹⁵ Définition proposée par Couperin <http://www.couperin.org/fr/groupe-de-travail-et-projets-deap/discovery/item/916-journee-detude-discovery-tools-du-8-septembre-2011>

disponibles sur le marché académique français et plus particulièrement aux éditeurs francophones, qui constituent la niche que les bases de connaissances internationales traitent le plus partiellement.

Cette base de connaissances est une composante du hub de métadonnées.

Le hub de métadonnées

La création d'un hub de métadonnées⁹⁶, axe à part entière du projet d'établissement de l'ABES, repose sur le constat qu'aujourd'hui les éditeurs de ressources électroniques produisent des métadonnées et les fournissent aux bibliothèques. La tendance était jusqu'à présent de chercher à négocier avec ces éditeurs la fourniture de données dans des formats adaptés aux formats des catalogues de bibliothèques (MARC) ce qui représentait un investissement non négligeable et pouvait avoir pour résultat des données relativement appauvries. Le hub repose sur l'idée de récupérer les données telles qu'elles sont produites par les éditeurs, dans leur format d'origine et donc dans toute leur richesse. Il s'agit ensuite de travailler sur ces données à l'état brut. Elles sont modélisées, corrigées, enrichies et peuvent être converties dans plusieurs formats (MARC, KBART⁹⁷, RDF⁹⁸).

Les corpus choisis en priorité sont ceux qui font l'objet des négociations dans le cadre des licences nationales, afin d'alimenter la plate-forme ISTEEX. Par ailleurs le hub est aussi un outil qui sert cette négociation en permettant de contrôler et d'affiner les données fournies par les éditeurs. Ainsi il est possible de comparer une liste de périodiques proposée par l'éditeur au registre ISSN, ce qui permet de révéler l'existence de fusions ou de changements de titres. L'ABES peut ainsi lors des négociations regrouper des titres en une seule et même archive.

Le hub permet également d'enrichir les listes de périodiques avec des métadonnées au niveau des articles, mêlant ainsi plusieurs niveaux de granularité. Ces données sont ensuite utilisées pour décrire le contenu d'un bouquet pour Bacon, pour signaler les ressources dans le Sudoc et bientôt pour la plate-forme ISTEEX. Une autre application de la gestion fine de ces métadonnées sur le long terme, les ressources électroniques étant acquises définitivement dans le cadre d'ISTEEX, est la possibilité pour les établissements de faire du désherbage des collections papier.

SGBm et Sudoc II

Un autre axe du projet d'établissement consiste à « proposer aux établissements un système de gestion mutualisé » (SGBm).

Actuellement de nombreux SIGB locaux coexistent dans le réseau, générant des coûts d'acquisition et de maintenance importants pour les établissements. L'idée qui sous-tend le projet de SGB (système de gestion de bibliothèque) mutualisé est de rechercher une mutualisation de ce travail et de ces coûts, dans le cadre de l'émergence d'une nouvelle offre, les SGB de nouvelle génération. Ceux-ci sont des offres hébergées, fonctionnant non plus par achat de licence mais sur une logique de services qui sont censés être évolutifs. Les SGB sont fournis avec des données bibliographiques et d'autorité et s'appuient sur une base de connaissances. Ce sont des systèmes modulables (via des API) qui permettent de répondre à l'inadaptation des SIGB traditionnels à la

⁹⁶ NICOLAS Yann, « Focus sur le hub de métadonnées », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canalc2.tv/video.asp?idvideo=12651>

⁹⁷ KBART (*Knowledge Bases And Related Tools*) est une recommandation d'origine britannique, qui va faire l'objet d'une norme NISO et qui a pour but de décrire les bouquets commerciaux de livres et de revues électroniques.

⁹⁸ RDF (*Resources description framework*) est le langage utilisé pour le web de données. Il en sera question au cours du second chapitre.

gestion d'une multiplicité de ressources (ressources papier, numériques, outils de découverte, archives ouvertes..)

Le recours à ce type d'offre suscite quelques réticences. Il s'agit d'une offre très récente, à dominante anglo-saxonne, sur laquelle l'on dispose encore de peu de recul. Le projet est donc mené prudemment, par étapes. Depuis septembre 2014, sept sites pilotes mènent un dialogue compétitif avec des éditeurs afin de sélectionner la meilleure offre, dans l'objectif de passer au nouveau système en 2016. Environ quarante établissements se sont déclarés prêts à signer l'accord-cadre dans l'hypothèse d'une réinformatisation de leur système survenant dans les trois ans.

Dans ce cadre, qu'advient-il du Sudoc ? Dans la mesure où différentes solutions pourront être retenues, car il s'agit d'un marché multi-attributaire, on pouvait supposer que le projet SGBm impliquerait la fin du catalogage partagé. D'autre part le logiciel OCLC qui est au cœur du Sudoc ne sera plus maintenu d'ici cinq ans. Il a cependant été décidé de traiter séparément les deux questions.

Un groupe de travail sur l'avenir du Sudoc a été formé⁹⁹, permettant de poser certains principes de base. Il a été acté que les membres du réseau souhaitent continuer à travailler dans un catalogue centralisé, qui serait cependant intégré dans les systèmes locaux (c'est-à-dire dont les mises à jour se feraient de façon fluide et non toutes les nuits). Il reste cependant encore beaucoup d'inconnues sur ce futur système. Par exemple la question du modèle de données n'a pas été tranchée. L'objectif affiché est de résoudre les difficultés posées par la coexistence des ressources physiques et électroniques. Le Sudoc II devra signaler la documentation papier et électronique de manière unifiée, ce qui permet à l'ABES d'afficher une certitude : le hub de métadonnées et Bacon en seront des modules à part entière.

Ces nouveaux projets de l'ABES font ressortir deux tendances majeures permettant de situer le positionnement stratégique de l'ABES.

La première est comme le pointe l'AERES, la « dimension d'agence de mutualisation au service des établissements dans laquelle l'ABES s'investit de plus en plus¹⁰⁰ ». La seconde tendance est la stratégie d'indépendance des données. Déjà, la stratégie des API visait à s'affranchir des outils propriétaires et, comme le montre l'application IdRef, à conserver la maîtrise sur les données. Ici, le projet Bacon vise à disposer de données fiables et à les partager, indépendamment des outils. Le hub cherche à rendre les données interopérables avec d'autres systèmes, y compris en dehors du réseau sur le web de données.

Ces projets nous rappellent que les métadonnées sont au cœur du métier de l'ABES. Raymond Bérard l'explique en ces termes : « Cette diversification est [...] cohérente avec les métiers de l'agence et ses compétences qui demeurent centrés sur une spécialité, le traitement des métadonnées¹⁰¹. »

⁹⁹ Groupe de préfiguration projet Sudoc II, « Sudoc II. Année 0 », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

¹⁰⁰ AERES, *op.cit.*

¹⁰¹ BÉRARD Raymond, *op.cit.*

LES DÉBUTS DE LA COOPÉRATION

L'ABES est née la même année que la BnF dans sa nouvelle structure. La naissance d'une seconde agence bibliographique a pu tout d'abord sembler d'une certaine manière concurrencer la BnF. Mais l'arrivée de l'ABES a surtout créé de fait un second public, celui des bibliothèques de l'enseignement supérieur, alors que la BnF était auparavant supposée s'adresser à tout type de bibliothèque. Les deux institutions se retrouvent alors chacune avec un périmètre respectif. Mais dans le sien, l'ABES est l'agence qui a réussi à mettre en place un réseau efficace, faisant peut-être figure de bonne élève. Peut-être est-ce une des raisons pour lesquelles les relations sont au début un peu difficiles entre les deux établissements. Des échanges de données sont mis en place dès le début, il existe une convention de coopération, mais celle-ci reste assez formelle et donne lieu à peu de projets concrets. Mais à partir de 2010 les relations s'intensifient et semblent prendre un nouveau départ.

Le signalement des ressources

Les échanges de notices

Les deux établissements viennent de renouveler la convention triennale qui les lie depuis 1998 et dont le premier objectif est de « définir les conditions et modalités d'échanges de notices ». En effet dès 1998, la BnF a fourni ses données d'autorité à l'ABES pour le chargement du Sudoc. Par la suite les deux établissements ont commencé à s'échanger les données suivantes :

La BnF fournit à l'ABES les notices bibliographiques et d'autorité de son catalogue général. Certaines de ces notices sont versées directement dans le Sudoc, d'autres dans une base d'appui qui vient en aide au catalogage. Il s'agit des notices des livres imprimés reçus par la BnF au titre du dépôt légal, des notices des périodiques du domaine français, et des notices d'autorités (personnes, collectivités et titres). La périodicité de ces envois est bimensuelle. On peut noter que cet échange pourrait aller dans les deux sens, le Sudoc pourrait en effet au même titre que Worldcat constituer pour la BnF un réservoir de notices à dériver, mais cela n'est pas le cas.

L'ABES fournit à la BnF les notices descriptives des centres de ressources participant au Sudoc, afin de constituer le Répertoire national des bibliothèques et centres de documentation (RNBFD). La périodicité de ces envois est mensuelle.

Ces échanges se font via le protocole FTP. Les données des établissements sont placées sous la licence ouverte de l'État qui impose la mention de provenance ou d'attribution, ce qui implique le maintien d'au moins l'identifiant pérenne de la notice. Mais la réutilisation, la diffusion, et la modification de ces notices sont autorisées.

Élaboration de Rameau

Rameau, le Répertoire d'autorité-matière encyclopédique et alphabétique unifié résulte en partie d'une collaboration entre la BnF et le réseau Sudoc même s'ils n'en sont pas les seules composantes. Des bibliothèques de lecture publique, de musées, de grandes institutions font aussi partie du réseau, de même que des bibliothèques étrangères francophones. Globalement les propositions émanent à 40 % de la BnF, à 40 % du Sudoc et à 20 % des autres contributeurs¹⁰².

¹⁰² Comité d'orientation Rameau, Compte-rendu de la réunion du 2 décembre 2013, Paris, Bibliothèque nationale de France, janvier 2014. Disponible sur http://rameau.bnf.fr/informations/pdf/comite_orientation_rameau_cr_20131202.pdf

Le Centre national Rameau (CNR) est rattaché à la BnF au département de l'IBN. Les propositions émanant du réseau doivent être approuvées par le CNR pour être prises en compte¹⁰³. Il s'agit donc d'un système collaboratif mais dont la direction est assurée par la BnF. Il s'agit à ce jour du seul exemple de fichier national d'autorités.

Le Catalogue Collectif de France

Géré par le département de la Coopération de la BnF, le CCFr est un catalogue collectif qui permet d'interroger simultanément un grand nombre de catalogues de bibliothèques françaises, dont ceux de la BnF et de l'ABES. Il s'agit du plus grand catalogue français. Il permet de localiser plus de 30 millions de documents, signalés dans les applications suivantes :

- Le catalogue général de la BnF.
- Le Sudoc.
- La base Patrimoine.
- Des catalogues de grandes bibliothèques municipales ainsi que le réseau des bibliothèques spécialisées de la ville de Paris.
- Le catalogue Archives et manuscrits de la BnF (BAM).
- Calames, le catalogue des archives et des manuscrits de l'enseignement supérieur.
- Le catalogue général des manuscrits des bibliothèques publiques de France (CGM).
- Le répertoire des manuscrits littéraires français du XX^{ème} siècle (PALME).

Le CCFr donne également accès au répertoire national des bibliothèques et des fonds documentaires (RNBFD) mentionné plus haut. Il offre en outre un service de prêt entre bibliothèques.

La mise en commun d'une réflexion stratégique

Au delà de ces réalisations très concrètes, les deux établissements ont à cœur d'avoir une réflexion commune sur les enjeux qui se présentent à eux.

Des réunions régulières ont lieu entre les deux établissements et leurs tutelles. Elles se tenaient auparavant dans un cadre nommé « SU-SI-CCFr » (SU désignant le Sudoc et SI le système d'information de la BnF), cadre qui était donc très largement orienté vers les outils. Il s'agissait en effet avant tout de points d'informations concernant les catalogues. En 2012 il est décidé d'adopter un nouveau nom pour le cadre de ces rencontres, le Comité stratégique bibliographique (CSB), dans l'objectif d'élaborer ensemble une vision stratégique nationale, qui prenne en compte, au-delà des catalogues, les problématiques de signalement en général. Les grands axes de cette stratégie sont exposés dans un rapport d'orientation diffusé en novembre 2012¹⁰⁴.

¹⁰³ Du point de vue du Sudoc l'intégration des autorités Rameau est plus fluide que pour les autorités qu'elle récupère de la BnF. Lorsque la mise à jour bi-mensuelle a lieu les vedettes matières modifiées sont automatiquement mises à jour dans le Sudoc.

¹⁰⁴ ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique préparé par l'Agence bibliographique de l'enseignement supérieur et la Bibliothèque nationale de France*, novembre 2012. Disponible sur

La dernière convention qui les lie, signée le 24 juin 2014, en plus de fixer les modalités d'échange de données, présente « les axes de coopération et d'expérimentation retenus pour les années à venir. » Ces sujets, sur lesquels nous aurons l'occasion de revenir, sont l'objet de réunions opérationnelles entre les équipes de l'ABES et du département de l'IBN.

La BnF et l'ABES se rencontrent également dans les instances de normalisation telles que le groupe sur l'adoption de RDA en France, dont il sera question un peu plus loin.

Des rôles bien définis

Nous voyons maintenant que les missions des deux établissements sont bien distinctes. La BnF, en charge du signalement et de la conservation du patrimoine national, est confrontée à des problématiques d'héritage, de transmission pour les générations à venir. A ce titre la conservation est pour elle un enjeu fort. Elle a aussi pour cœur de métier la gestion de collections spécialisées (monnaies et médailles, estampes, photographies, cartes et plans, musique, chorégraphies, documents sonores...). L'ABES, responsable du signalement des thèses et des ressources électroniques, se positionne davantage sur des problématiques d'accès, même si son rôle dans les licences nationales touche aussi à la conservation pérenne, celle-ci étant assurée par le CINES¹⁰⁵ (Centre Informatique National de l'Enseignement Supérieur). Par ailleurs elle est confrontée à davantage de documentation étrangère.

Les deux institutions renvoient également des images bien différenciées, qui sont le reflet de cultures d'entreprise différentes. En témoigne l'hommage que rendait le directeur de l'IBN à l'ABES pour ses 20 ans : « Au regard des siècles qui pèsent sur la vieille dame de la rue de Richelieu et face au vertige un peu écrasant des tours de verre de la citadelle de Tolbiac, l'ABES donne l'image d'une entreprise à taille humaine, presque une *start-up* californienne, à la pointe des questions nouvelles et compliquées¹⁰⁶. »

Il est vrai que l'ABES est une agence, et à ce titre elle est dotée d'une certaine agilité. Raymond Bérard, en évoquant le rapport du Conseil d'État sur les agences¹⁰⁷, rappelle que leur force réside « dans leur capacité à se focaliser sur une spécialité, à savoir se doter des expertises requises et à réagir rapidement en raison de leur autonomie de gestion¹⁰⁸ ». La BnF à l'inverse peut parfois faire preuve d'une certaine lourdeur. Même si face à l'ABES, l'IBN qui est un département de taille comparable, est lui aussi doté d'experts « à la pointe des questions nouvelles », la structure hiérarchique de la BnF, la longueur des circuits de décision, rendent celle-ci beaucoup moins souple et agile.

Nous l'avons évoqué, les deux établissements sont distincts par la nature de leur réseau. La BnF n'a pas un réseau structuré comme celui de l'ABES. La décentralisation en 1982 et 1983 a rendu difficile la constitution d'un réseau de catalogage partagé pour les bibliothèques publiques. Mais la BnF est néanmoins en train de construire d'autres

http://www.bnf.fr/documents/rapport_csb.pdf

¹⁰⁵ <https://www.cines.fr/>

¹⁰⁶ ILLIEN Gildas, « 20 ans, le bel âge ! », in *Arabesques*, n°75, juillet 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-75>

¹⁰⁷ Conseil d'État, *Rapport public 2012 du Conseil d'État - Les agences : une nouvelle gestion publique ?*, Paris, La Documentation française, 2012. Disponible sur <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/124000501/0000.pdf>

¹⁰⁸ BÉRARD Raymond, *op.cit.*

types de réseaux (pôles associés, Gallica marque blanche...). Par ailleurs elle se trouve par son statut de bibliothèque nationale confrontée à des utilisateurs de tout type qui se situent bien au-delà des frontières du territoire français. Cette grande visibilité peut ainsi justifier le maintien d'une culture monolithique, qui se traduit par une certaine rigueur, une imperméabilité à ce qui n'est pas produit par elle, un jargon très touffu et une identité forte.

A l'inverse l'ABES a sans doute bénéficié de l'absence d'autonomie des universités jusqu'à la loi LRU en 2007, rendant l'adhésion des bibliothèques universitaires au Sudoc plus ou moins inévitable. L'autonomie des universités et la nécessité de maîtriser leurs ressources font que le réseau est d'autant plus aujourd'hui perçu par les établissements comme un levier de mutualisation. Mais ceci ne sera valable que tant que les services de l'ABES resteront performants et gratuits. Elle doit aussi parvenir à responsabiliser de plus en plus les établissements, les encourager à prendre des initiatives et valoriser celles-ci. Cela représente pour elle un enjeu de taille, car l'ABES véhicule aussi l'image d'un établissement un peu débordé, qui fait trop de choses en même temps. Son agilité lui permet de faire face à cette situation, mais il serait préférable qu'elle parvienne à une organisation plus stable, d'autant plus que la part d'emplois contractuels au sein de l'agence est très forte, ce qui pose la question de la sécurisation de ses missions¹⁰⁹.

Mais la transition bibliographique va quoi qu'il en soit redessiner le rôle des agences et renforcer leurs fonctions de mutualisation et de têtes de réseau.

¹⁰⁹ Rapport AERES : « les nouvelles missions de l'Abes ont été prises en charge par des agents contractuels (17 contractuels pour 73 agents), sur des crédits alloués par l'État, pour assurer de nouvelles activités. Ceci pose la question de la pérennisation des nouvelles missions de l'Abes.»

LA TRANSITION BIBLIOGRAPHIQUE

L'expression de « transition bibliographique¹¹⁰ » désigne le mouvement qui consiste pour les bibliothèques à s'adapter à une mutation profonde des modes d'accès à l'information, celle-ci étant liée principalement à l'arrivée du numérique et du web. Nous en avons eu un aperçu en décrivant les missions des deux agences bibliographiques. L'adaptation du dépôt légal pour la BnF, la création d'un hub de métadonnées pour l'ABES sont des projets qui entrent dans le cadre de cette mutation. Mais plus radicalement les bibliothèques doivent revoir tout leur dispositif de production de métadonnées bibliographiques.

Nous proposons d'effectuer une description de ces changements induits par le numérique et le web, puis nous verrons comment la sphère des bibliothèques s'adapte en faisant évoluer ses standards et outils avant de nous demander comment ces mutations touchent aussi profondément les façons de travailler.

L'ÉVOLUTION DES COLLECTIONS, DES BESOINS ET DES USAGES

Les ressources numériques changent la donne

Les ressources électroniques comportent des spécificités nouvelles. Avec le numérique la notion de reproduction en tant que telle disparaît tandis que de nouveaux modes de diffusion ou de mise à jour apparaissent (chapitre par chapitre par exemple). Une ressource peut avoir une existence sous plusieurs formes et être diffusée simultanément selon différents modes. Ainsi un ouvrage de la collection *Que-sais-je ?* sous sa forme numérique peut être commercialisé à l'unité ou dans le cadre d'un bouquet comme le propose l'agrégateur Cairn.

Par ailleurs des niveaux de granularité différents coexistent, rendant le signalement complexe. Le rapport pour le Comité stratégique bibliographique souligne ce point : « la recherche requiert un traitement à un niveau de granularité toujours plus fin », au niveau de l'article, des sommaires, des chapitres, des pages, des mots. En même temps, « il sera tout autant nécessaire de se repérer et de naviguer dans de très grandes masses de données brutes, fonds, recueils ou corpus numériques¹¹¹ ». Il existe en outre des types de ressources inédits à prendre en compte, tels que les jeux de données de la recherche¹¹². De plus en plus les chercheurs vont être incités à gérer les données qu'ils produisent de manière à permettre leur diffusion et leur réutilisation, ce qui va nécessiter de prendre en compte ces nouveaux corpus, en complément des thèses et des archives ouvertes.

Les bibliothèques doivent donc développer des modes de signalement adaptés à cette coexistence de niveaux de granularité, proposer des services de découverte (dans lesquels une seule recherche permet d'accéder à tout type de ressource) ou des services complémentaires tels que le *data mining*. En effet « dans le monde de

¹¹⁰ L'expression a été forgée à la BnF et inspirée par l'initiative de la Bibliothèque du Congrès BIBFRAME qui signifiait au départ *Bibliographic Framework Transition Initiative*. Aujourd'hui le terme *transition* a disparu de l'appellation officielle qui est devenue *Bibliographic Framework Initiative*. <http://www.loc.gov/bibframe/>

¹¹¹ ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique*, *op.cit.*

¹¹² Sur ce sujet on consultera avec profit le mémoire de Rémi GAILLARD, *De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?*, mémoire DCB sous la direction de Monique Joly, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/notices/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche>

la recherche on ne « lit » plus forcément les documents signalés au catalogue, on a aussi besoin de « traiter » et de « fouiller » des corpus de métadonnées pour analyser des séries, comparer des résultats issus de sources différentes, calculer la notoriété des références, des auteurs¹¹³ ».

Autre spécificité des ressources électroniques, alors que la question centrale qui se pose lorsque l'on gère des documents imprimés est celle du stock, celle qui survient avec le numérique est celle de l'accessibilité, avec des modèles fondés non plus sur la possession mais sur l'abonnement. Ce type de modèle pose particulièrement problème à la BnF, dont la mission repose sur la transmission du patrimoine aux générations futures.

Parallèlement, dans un contexte de crise économique, les bibliothèques sont soumises comme tous les services publics à des budgets contraints, tout en devant faire face à une augmentation des ressources à signaler. Ces contraintes budgétaires incitent les bibliothèques à rechercher des pistes pour faire des économies. Il devient capital qu'elles se concentrent sur leur valeur ajoutée et collaborent entre elles mais aussi avec d'autres acteurs du monde culturel et scientifique. On retrouve cette idée dans le projet d'établissement de l'ABES : « Dans un contexte où les ressources que les bibliothèques mettent à disposition de leur public excèdent de très loin les seules ressources imprimées, le catalogue doit être repensé en prenant acte de cette évolution et en le positionnant dans une vision beaucoup plus globale de l'accès à l'information et aux ressources¹¹⁴ ». Cette vision plus globale dans laquelle inscrire les catalogues est celle du web.

Le web change la donne

Avant l'arrivée du numérique, la notion de publication était centrale. Le modèle du mode d'accès à l'information passait par un objet imprimé, par un document (article, revue, monographie). Dans ce modèle les bibliothèques avaient une forte légitimité, elles étaient perçues comme incontournables pour l'accès au savoir, de même que les catalogues qui en étaient les principaux points d'accès.

Avec le numérique et le web, l'information est devenue disponible immédiatement et en abondance. La quantité de ressources existantes s'accroît de manière exponentielle, et l'accès au texte intégral tend à devenir la norme. L'internaute peut donc accéder directement à cette masse d'informations sans l'intermédiaire des bibliothèques. Cette situation met fin à l'idée que les bibliothèques détiennent le monopole de l'accès au savoir. Cette opinion n'est pas peu répandue puisque même le ministre de la Culture avait déclaré en 2006 : « avec Internet ce n'est plus la peine d'aller à la bibliothèque¹¹⁵ ».

Les habitudes des utilisateurs en ont été profondément modifiées. La recherche sur le web n'a plus grand-chose à voir avec la recherche dans des systèmes documentaires. Ces derniers sont généralement caractérisés par une homogénéité des sources et des informations, recouvrant des domaines bien circonscrits, des bases structurées, une indexation avec un langage contrôlé adapté au domaine, des outils d'interrogation uniques. Tandis que le web rassemble des

¹¹³ ILLIEN Gildas, « Décrire les objets du savoir, les nouveaux paradigmes du catalogage », in *Documentaliste-Sciences de l'Information*, vol.50, n°3, octobre 2013, p.26-29

¹¹⁴ ABES, *Projet d'établissement 2012-2015*, novembre 2011. Disponible sur <http://www.abes.fr/Connaitre-l-ABES/Projet-d-etablissement>

¹¹⁵ Renaud Donnedieu de Vabres, France 2, 3 mars 2006

sources et des approches hétérogènes (scientifique, commerciale, vulgarisation...), diffusées le plus souvent sur des sites non structurés, sans langage d'indexation et qu'il mêle différents modes de recherche (répertoires, saisie de termes dans les moteurs de recherche, navigation de lien en lien).¹¹⁶

Comme l'explique par ailleurs Emmanuelle Bermès : « Les millions de résultats renvoyés par Google sur n'importe quelle recherche, assortis avec son classement des résultats par pertinence d'une performance jusque-là inégalée, révoquent les pratiques les plus fondamentales de l'usage traditionnel des catalogues, comme l'accès chronologique ou alphabétique. Ce qu'on pourrait considérer comme un détail n'en est pas un lorsqu'on réalise à quel point les pratiques des usagers ont été bouleversées par ce changement. Un internaute qui ne trouve pas la réponse qu'il recherchait dans la première page de résultats du moteur va formuler une nouvelle recherche plutôt que de parcourir la deuxième page. »¹¹⁷

Les attentes des usagers sont donc modifiées en termes de qualité et de pertinence des réponses. Les moteurs de recherche répondent à ces attentes, (ou les créent plus probablement) en proposant désormais une mise en contexte des résultats. Ainsi Google explique en 2012 dans un billet de blog intitulé *Introducing the Knowledge Graph: things, not strings*¹¹⁸ que les internautes attendent désormais dans les résultats des *entities* (entités : personnes, lieux, sujets...), comprenons des résultats mis en contexte en fonction de leur statut et non plus simplement des chaînes de caractère (*strings*). Cela se traduit par des fonctions permettant de désambiguïser une recherche (pour déterminer par exemple si une requête « Taj Mahal » désigne le monument ou un restaurant indien), d'offrir un résumé des informations principales dans la liste des résultats sous la forme d'une *knowledge card* (pour un auteur on pourra voir en un coup d'œil son portrait, ses dates de naissance et de mort, ses œuvres principales...) et de suggérer des recherches associées.

Les moteurs de recherche sont donc désormais le point d'entrée à toute recherche d'information. Comme le préconise le rapport d'orientation pour le CSB, « cette nouvelle donne implique de reformuler l'offre bibliographique publique selon les usages et l'environnement des internautes plutôt qu'en fonction des contraintes et exigences de production des bibliothécaires, encore très prégnantes dans la conception des catalogues¹¹⁹ ». Or les données des catalogues de bibliothèques sont aujourd'hui cantonnées dans ce que l'on appelle le « web profond » ou « web invisible ». Elles ne sont en effet accessibles que via des formulaires de recherche, barrières que ne franchissent pas les robots d'indexation des moteurs de recherche.

L'enjeu pour les bibliothèques est donc de parvenir à rendre visibles leurs données dans la masse d'information disponible sur le web. Cet enjeu est de taille aussi pour les internautes car les bibliothèques produisent une information bibliographique de grande qualité, validée, normalisée et dans laquelle on peut raisonnablement avoir confiance. Dans le cadre d'une recherche d'information sur un créateur, exemple assez banal et fréquent, il est évident que les bibliothèques disposent d'informations à forte valeur ajoutée, à travers sa notice d'autorité, elle-même riche de renseignements, ou encore la liste de ses œuvres et de ses différentes contributions et leurs caractéristiques. Il est tout à fait intéressant pour l'utilisateur du web que ces informations puissent être affichées dans les résultats d'une recherche.

¹¹⁶ MKADMI Abderrazak, SALEH Imad, *Bibliothèque numérique et recherche d'informations*, Paris, Hermès Science, Lavoisier, 2008, p.56

¹¹⁷ BERMES Emmanuelle, avec la collaboration d'Antoine ISAAC et Gautier POUPEAU, *Le Web sémantique en bibliothèque*, Paris, Ed. du Cercle de la Librairie, 2013

¹¹⁸ <http://googleblog.blogspot.fr/2012/05/introducing-knowledge-graph-things-not.html>

¹¹⁹ ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique, op.cit.*

Les bibliothèques portent donc un intérêt croissant aux technologies du web sémantique, qui leur permettraient de diffuser leurs données au-delà des frontières du catalogue. En témoigne le développement d'une offre de formation à destination du personnel des bibliothèques et dont un intitulé revient avec régularité, « les catalogues au défi du web¹²⁰ ».

Bibliothèques et web de données

Revenons brièvement sur les concepts fondamentaux du web de données avant de nous pencher sur l'intérêt que lui portent les bibliothèques, intérêt qui s'est notamment manifesté par la création du groupe de travail du W3C¹²¹ « Bibliothèques et web de données » et qui a donné lieu à un rapport publié en 2011¹²².

Lorsque Tim Berners-Lee invente le web au début des années 90, il est à la recherche d'un moyen de permettre aux chercheurs du CERN de partager leurs documents stockés sur des ordinateurs distants les uns des autres. Le web, qui est donc un réseau de documents (à distinguer d'Internet qui désigne le réseau technique, le réseau des machines), repose sur 4 composantes :

- Le principe de l'hypertexte, qui permet de lier les pages entre elles.
- Le protocole HTTP (*Hypertext Transfer Protocol*) qui permet aux machines de communiquer.
- Un mécanisme d'identification des ressources, URL (*Uniform Resource Locator*) ou URI (*Uniform Resource Identifier*) qui permet sans connaître la machine où se trouve la ressource de pointer vers elle. (Une URL est une catégorie d'URI utilisant le protocole HTTP).
- Un langage de description des pages web, HTML (*Hypertext Markup Language*).

On parle donc pour ce premier web de « web de documents ». Mais Tim Berners-Lee, dans la continuité de ce projet, suggère de passer à un « web de données¹²³ ». Il s'agit de lier non plus les documents entre eux mais directement les données. Seuls les humains sont vraiment en mesure de comprendre les documents, en les lisant, tandis que l'exploitation des données offre de nombreuses possibilités et surtout peut aussi être menée à bien par des machines.

Cela suppose de pouvoir identifier les données de manière fiable et de les lier entre elles avec des liens qui ont un sens. Le terme de web « sémantique » vient de l'idée qu'en typant ces liens on permet « à des machines d'interpréter la nature de la relation¹²⁴ » entre deux ressources.

Tim Berners-Lee publie en 2006 un article dans lequel il expose les quatre grands principes du web de données¹²⁵ :

¹²⁰ En octobre 2014 une formation de 3 jours intitulée ainsi a été dispensée à l'ENSSIB dans le cadre de la formation continue. En novembre, c'est une journée de sensibilisation à destination des cadres sur le même thème qui a été proposée par la BnF et le CNFPT.

¹²¹ Le W3C (*World Wide Web consortium*) est l'organisme de normalisation du web. <http://www.w3.org/>

¹²² GROUPE D'INCUBATION W3C, *op.cit.*

¹²³ BERNERS-LEE Tim, « Le web à venir », [enregistrement vidéo], in TED Conférences, février 2009. Disponible sur http://www.ted.com/talks/tim_berniers_lee_on_the_next_web?language=fr

¹²⁴ BERMES Emmanuelle, *op.cit.*

¹²⁵ BERNERS-LEE Tim, *Linked Data*, 2006. Disponible sur <http://www.w3.org/DesignIssues/LinkedData.html>

- Utiliser des URI pour identifier des ressources.
- Formuler ces URI selon le protocole HTTP.
- Lier les ressources entre elles.
- Recourir aux standards RDF et SPARQL (*SPARQL Protocol and RDF Query Language*). Ces deux standards sont développés et maintenus par le W3C.

SPARQL est un protocole et un langage de requête, comparable à SQL (*Structured Query Language*) mais qui est adapté au formalisme RDF. Celui-ci repose sur une logique consistant à décrire la nature de la relation entre deux ressources sous la forme d'une phrase simple, qui suit la structure sujet – prédicat – objet et qui est appelée triplet.

Le sujet du triplet est obligatoirement identifié par une URI. Il en est de même pour le prédicat, grâce à l'existence de vocabulaires qui expriment les relations entre les ressources. Nous prenons ci-après un exemple tiré du *Dublin Core Metadata Terms*¹²⁶. L'objet peut également être une URI ou bien une simple chaîne de caractère (appelée un littéral).

Ainsi pour exprimer la phrase « *Northanger Abbey* a pour auteur Jane Austen », on peut employer le triplet RDF suivant.

<<http://viaf.org/viaf/175982694>><<http://purl.org/dc/terms/creator>><« Jane Austen »>

Si l'on souhaite remplacer le littéral « Jane Austen » par une URI, le triplet devient :

<<http://viaf.org/viaf/175982694>><<http://purl.org/dc/terms/creator>><<http://viaf.org/viaf/102333412>>

L'intérêt d'utiliser des URI au lieu de littéraux permet de continuer à créer des liens entre différents triplets, ce qui constitue un graphe de données. Comme l'indique le rapport du W3C : « dans un écosystème à base de graphes, une institution peut fournir un certain nombre de déclarations sur une ressource; toutes les déclarations ainsi fournies sur une ressource donnée, identifiée de manière unique, peuvent alors être agrégées en un graphe global¹²⁷. »

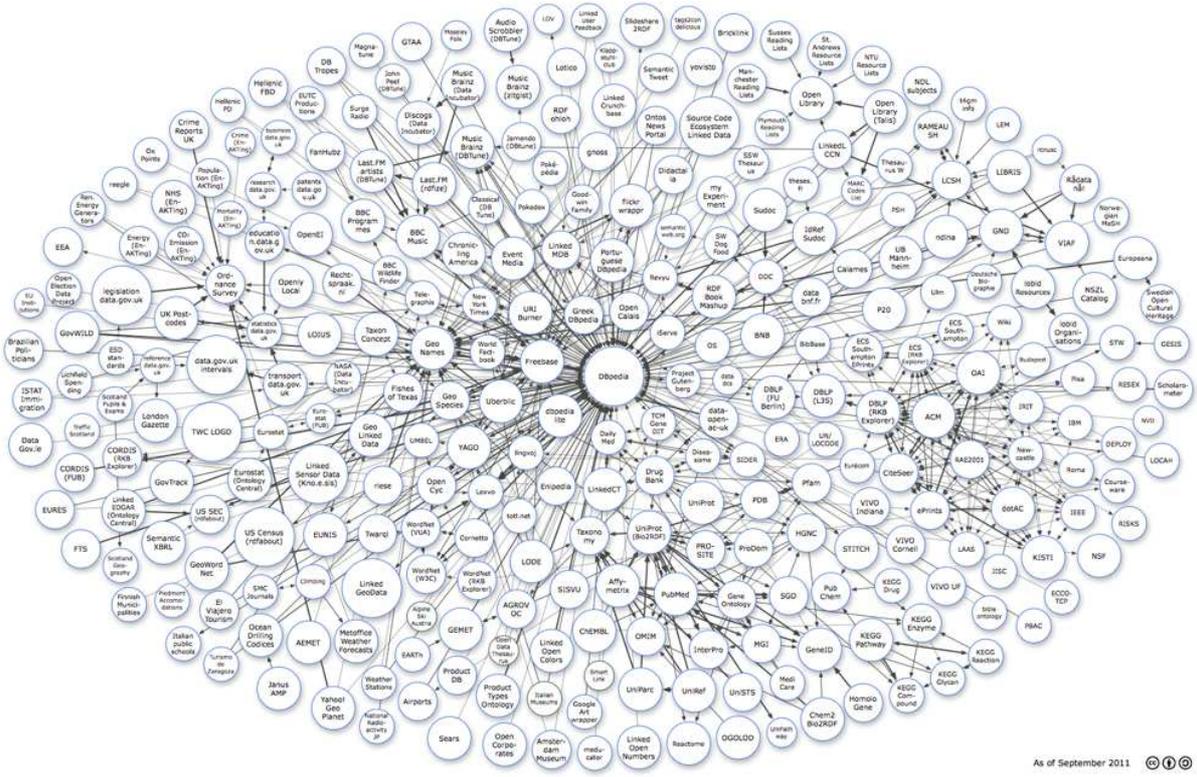
Ainsi le web de données permet à différents acteurs de mettre en commun leurs efforts. Cette approche concerne tous les domaines de la connaissance et permet de disposer aujourd'hui d'un immense graphe d'informations agrégeant un grand nombre de corpus de métadonnées. L'un des premiers jeux de données à alimenter ce corpus est DBpedia¹²⁸, issu d'une initiative de l'université de Berlin qui publie en 2007 les données contenues dans les pages de Wikipédia. DBpedia est aujourd'hui au cœur d'un graphe d'informations représenté sous la forme d'un nuage, le *Linking Open Data cloud*¹²⁹ :

¹²⁶ <http://dublincore.org/documents/2012/06/14/dcmi-terms/>

¹²⁷ GROUPE D'INCUBATION W3C, *op.cit.*

¹²⁸ <http://dbpedia.org>

¹²⁹ Richard Cyganiak, <http://lod-cloud.net/state/>. Il s'agit de l'état du nuage du LOD en septembre 2011, les évolutions plus récentes rendant le schéma illisible sur une page imprimée. Une version actualisée est consultable en ligne dans un format permettant de zoomer sur des zones du nuage, sur <http://lod-cloud.net/>.



As of September 2011

Parmi les premiers jeux de données liés à DBpedia se trouvent Geonames¹³⁰, une base de données géographiques comportant notamment des coordonnées permettant la géolocalisation ou encore FOAF¹³¹ (*Friend Of A Friend*), un vocabulaire développé dans le cadre des réseaux sociaux et qui permet de décrire les relations entre personnes. Lorsque l'on zoome dans le nuage sur le secteur « publications » on retrouve des jeux de données du monde culturel et des bibliothèques : Worldcat, Europeana, la Bibliographie nationale britannique, le Thésaurus W des Archives nationales de France, ainsi que les jeux de données de l'ABES (IdRef, theses.fr, Sudoc) et de la BnF (data.bnf) sur lesquels nous reviendrons.

Le web de données nécessite pour la description des relations entre entités des vocabulaires plus ou moins spécialisés selon les domaines. Nous avons évoqué le *Dublin Core Metadata Terms*, un vocabulaire assez générique permettant de décrire les ressources bibliographiques. Citons également SKOS¹³² (*Simple Knowledge Organisation System*) qui permet d'exprimer des concepts ou de décrire des thésaurus. Il a été utilisé pour traduire les langages d'indexation Rameau et LCSH (*Library of Congress Subject Headings*) que l'on trouve également dans le nuage du LOD.

L'adoption des standards du web sémantique permet de donner de la visibilité aux données des bibliothèques, mais cet objectif requiert aussi de s'intéresser aux stratégies de référencement des moteurs de recherche. Schema.org¹³³ est une initiative lancée par des moteurs de recherche (Bing, Google, Yahoo) en 2011 qui repose sur l'intégration dans les pages web d'un schéma de micro-données

¹³⁰ <http://www.geonames.org/>
¹³¹ <http://www.foaf-project.org/>
¹³² <http://www.w3.org/2004/02/skos/>
¹³³ <https://schema.org/>



permettant d'améliorer la présentation des résultats en y faisant figurer des informations pertinentes pour l'utilisateur.

A la suite de cela, un membre d'OCLC a lancé dans le cadre d'un groupe de travail du W3C une initiative complémentaire pour créer Bib Extend¹³⁴, un schéma visant à enrichir Schema.org là où celui-ci s'avérait insuffisant pour décrire les ressources bibliographiques. Il s'agit donc avec Bib Extend d'améliorer la représentation de l'information bibliographique dans les résultats d'une recherche sur le web.

Les bibliothèques mènent des efforts importants dans le cadre du W3C pour adopter les standards du web sémantique et rendre leurs données présentes sur le web. Mais cet effort passe aussi par la révision de leurs outils de travail traditionnels. C'est ce à quoi elles s'emploient depuis les années 90 : repenser leurs catalogues pour les adapter aux nouveaux usages de recherche d'information.

VERS UN NOUVEAU MODÈLE POUR LA DESCRIPTION BIBLIOGRAPHIQUE

Il est nécessaire de rappeler dans cette section quels sont les principes de catalogage et formats d'encodage avec lesquels les bibliothèques travaillent actuellement, avant d'esquisser les contours de la transition bibliographique en cours, dont l'enjeu consiste à « revoir radicalement un dispositif vieux de près de 40 ans sans pour autant perdre la richesse et la cohérence des pratiques de description qui font la valeur ajoutée des données de bibliothèques¹³⁵ ». Cette transition passe par la mise en œuvre du modèle de données FRBR, l'adoption du code de catalogage RDA et la réflexion autour de la révision des formats MARC, avec un objectif en toile de fond qui est la nécessité d'être présent sur le web de données.

Nos outils actuels

Les règles de description bibliographique

L'harmonisation des pratiques de catalogage sur le plan international se fait dès les années soixante. En 1961 les *Principes internationaux de catalogage* auxquels on se réfère sous le nom de *Principes de Paris*, sont promulgués sous l'égide de l'IFLA et de l'UNESCO.

Réunie en 1969 par l'IFLA, la Conférence de Copenhague aboutit en 1971 à la publication de la *Description bibliographique internationale normalisée* (ISBD, *International Standard Bibliographic Description*). Ce document précise les éléments requis pour une description bibliographique complète, indique leur ordre de présentation, la ponctuation qui les délimite et donne des règles pour leur transcription¹³⁶.

La première publication de l'ISBD en 1971 concerne le description de monographies : ISBD(M). Elle est suivi d'une déclinaison pour les publications en série en 1974 : ISBD(S) qui deviendra en 2002 ISBD(CR) pour les ressources continues et qui

¹³⁴ <http://www.w3.org/community/schemabibex/>

¹³⁵ ABES, BnF, Rapport d'orientation pour le Comité stratégique bibliographique, *op.cit.*

¹³⁶ BnF. Pour les professionnels. ISBD (International Standard Bibliographic Description) [en ligne]. Paris : Bibliothèque nationale de France. Disponible sur http://www.bnf.fr/fr/professionnels/normes_catalogage/a.normes_isbd_presentation.html

inclut les sites web. D'autres déclinaisons concernent les documents cartographiques : ISBD(CM) ; les « non-livres » : ISBD(NBM) qui prend en considération l'apparition de nouveaux supports (microformes, documents audiovisuels...) ; la musique imprimée : ISBD(PM) ; les livres anciens : ISBD(A) ; les ressources électroniques : ISBD(ER).

En 2007 une première version intégrée rassemble les textes de ces ISBD spécialisés. La version actuelle de référence est la version intégrée publiée en 2011¹³⁷. Une pause dans l'évolution des ISBD a été pratiquée dans les années 90 afin de mener une réflexion de fond qui débouchera sur les FRBR.

Les formats d'encodage

Le format MARC (*MACHine READable Cataloging*, que l'on peut traduire par catalogage informatisé) a été créé dans le courant des années soixante en tenant compte des travaux des *Principes de Paris*. Il s'agit en réalité d'une famille de formats, conçus dans le but d'échanger des notices bibliographiques, et qui sont très peu répandus en-dehors du monde des bibliothèques. Ils s'appuient sur les ISBD et évoluent en tenant compte de leur révision¹³⁸.

Chaque pays a développé son propre format MARC, situation à laquelle on se réfère sous le nom de « Babel MARC ». Les plus répandus aujourd'hui sont MARC21, maintenu par la Bibliothèque du Congrès, et l'UNIMARC, élaboré sous l'égide de l'IFLA et largement utilisé en France. La BnF a développé son propre format, INTERMARC, prenant en compte ses spécificités de travail.

La norme ISO 2709 définit un protocole permettant aux bibliothèques d'importer ou exporter des notices. Elle permet de réaliser des conversions entre les différents formats MARC.

Une nouvelle façon de modéliser l'information : FRBR

Qu'est-ce que le modèle FRBR ?

En 1990, le séminaire de Stockholm sur les notices bibliographiques, organisé par l'IFLA, a la volonté de prendre en compte l'évolution de l'information et propose de travailler à la redéfinition du contenu d'une notice bibliographique, en poursuivant deux objectifs :

- Définir un cadre permettant une adéquation entre les données qui figurent dans les notices bibliographiques et les besoins des utilisateurs. Il s'agit donc d'un modèle conceptuel développé sur la base des besoins des utilisateurs, qui sont les usagers des bibliothèques mais aussi les bibliothécaires, les éditeurs, les distributeurs ou encore les gestionnaires des droits de propriété intellectuelle.
- Préconiser un niveau minimal des fonctions devant être assurées par les notices produites par les agences bibliographiques nationales. Cette demande émanait des agences bibliographiques nationales qui devant

¹³⁷ IFLA, ISBD : Description bibliographique internationale normalisée, édition intégrée, Paris : Bibliothèque nationale de France, 2013. Disponible sur http://www.bnf.fr/documents/isbd_trad_francais.pdf

¹³⁸ LE PAPE Philippe, avec la collaboration de Françoise LERESCHE, « Le signalement des documents. Catalogues informatisés et communication de données », in SANZ Pascal, *Guide de la coopération entre bibliothèques*. Paris, Éd. du Cercle de la librairie, 2008.

l'accroissement des ressources à décrire souhaitaient disposer d'un modèle de notice minimal défini au niveau international (*Core Record*).

Ce processus aboutit en 1998 à la publication des FRBR (*Functional Requirements for Bibliographic Records*) ou *Fonctionnalités requises des notices bibliographiques*¹³⁹.

Dans le modèle FRBR, les entités qui sont les « objets fondamentaux pertinents pour les utilisateurs de données bibliographiques¹⁴⁰ », sont réparties en trois groupes.

Les entités du groupe 1 représentent « les produits d'une activité intellectuelle ou artistique ». Elles comprennent :

- L'œuvre, qui est une notion abstraite, une création. Reprenons notre exemple avec *Northanger Abbey* de Jane Austen.
- L'expression, qui est une réalisation de l'œuvre. Par exemple la traduction de *Northanger Abbey* en français par Pierre Arnaud.
- La manifestation, qui est une concrétisation de l'expression de l'œuvre. Par exemple l'édition de *L'Abbaye de Northanger* par Gallimard en 2004 dans la collection *L'imaginaire*. Notons que c'est au niveau de la manifestation que sont actuellement constitués nos catalogues.
- L'item, qui est un exemplaire donné de la manifestation, par exemple celui que conserve la Bibliothèque municipale de Lyon.

Les entités du groupe 2 représentent les personnes physiques ou morales qui ont une responsabilité dans la création des entités du groupe 1 (création, réalisation, production, possession). Elles comprennent :

- Les personnes.
- Les collectivités.

Dans notre exemple, Jane Austen est à l'origine de la création de l'œuvre, Pierre Arnaud a la responsabilité de la réalisation de l'expression, Gallimard celle de la production de la manifestation, la Bibliothèque Municipale de Lyon celle de la conservation d'un item.

Les entités du groupe 3 traitent du sujet des œuvres, qui peuvent être des concepts, des objets, des événements ou des lieux.

Chacune de ces entités des trois groupes comporte des attributs permettant de la qualifier. Ainsi les attributs de l'œuvre sont notamment son titre, sa forme, son contexte ou encore le public visé.

Le modèle FRBR définit également des relations entre ces entités, d'où sa qualification de modèle « entité relation ».

Le modèle FRBR étant centré sur les notices bibliographiques, il a ensuite été étendu aux autres informations présentes dans un catalogue de bibliothèque :

- Les données d'autorité avec le modèle FRAD en 2009 (*Functional Requirements for Authority Data*)¹⁴¹. Ce modèle traite des entités « personne », « collectivité », « famille » et « œuvre ».

¹³⁹ Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, *Fonctionnalités requises des notices bibliographiques : rapport final*, Paris, Bibliothèque nationale de France, 2012. Disponible sur http://www.bnf.fr/documents/frbr_rapport_final.pdf

¹⁴⁰ *op.cit.*

¹⁴¹ Groupe de travail IFLA sur les Fonctionnalités requises et la numérotation des notices d'autorité (FRANAR), *Fonctionnalités requises des données d'autorité : un modèle conceptuel : rapport final*, Paris, Bibliothèque nationale de France, 2010. Disponible sur http://www.bnf.fr/documents/frad_rapport_final.pdf

- Les données d'autorité matière en 2010 avec le modèle FR SAD (*Functional Requirements for Subject Authority Data*)¹⁴².

Des travaux sont actuellement menés dans le cadre de l'IFLA sur la consolidation des modèles FRBR, FRAD et FR SAD.

Quels sont les avantages des FRBR ?

Le modèle FRBR opère un changement relativement important dans la conception des catalogues, en permettant de passer d'une prépondérance du document (contenant) à l'œuvre (contenu). En effet un catalogue traditionnel a pour principale fonction la description et la localisation des supports présents dans une bibliothèque. Dans le modèle préconisé par les ISBD, le catalogage se pratique d'ailleurs avec le document en main, de manière à pouvoir transcrire les informations telles qu'elles figurent sur le document. Un catalogue « FRBRisé » (un barbarisme qu'il est difficile d'éviter lorsque l'on aborde ces sujets) est avant tout un catalogue d'œuvres, ce qui constitue une approche beaucoup plus intuitive pour l'utilisateur.

On présente également souvent le modèle FRBR comme favorisant la sérendipité. En plaçant les œuvres dans un réseau de liens à parcourir il permet de faciliter les découvertes fortuites.

Un autre avantage de ce modèle, qui décortique les entités et leurs relations, est qu'il ouvre la voie à la mise à disposition des informations de la notice sous une forme éclatée. « Grâce à l'analyse proposée par FRBR, la notice descriptive cesse d'être perçue comme un bloc d'information compact pour s'apparenter plutôt à un réseau d'éléments de données ayant entre eux de multiples relations¹⁴³. » On entrevoit déjà combien cette nouvelle modélisation est adaptée à la dissémination sur le web des informations bibliographiques.

Le rôle central des autorités

Le modèle FRBR est aussi un modèle qui met les autorités au premier plan. Dans le catalogage traditionnel, il s'agissait de partir d'un document pour créer des points d'accès à ce document, les vedettes ou autorités. La démarche du modèle FRBR est inverse, c'est l'autorité qui est l'élément central auquel on lie les documents.

Rappelons brièvement les fonctions d'une notice d'autorité :

- Normaliser la forme des points d'accès, c'est-à-dire indiquer la forme retenue et les formes rejetées. La prise en compte de formes rejetées permet à l'utilisateur de retrouver la notice quelle que soit la clé de recherche dont il se sert.
- Identifier l'entité sans ambiguïté, via l'attribution d'un identifiant.
- Apporter des informations sur l'entité (justification du choix de la forme retenue, notes sur les sources consultées pour établir la notice).

Les fichiers d'autorités servant à apporter une aide lors du catalogage, ils sont à l'origine liés à un catalogue donné. Cependant ils ont tendance à devenir de plus

¹⁴² Groupe de travail IFLA sur les Fonctionnalités requises des notices d'autorité matière (FRSAR), *Fonctionnalités requises des données d'autorité matière : un modèle conceptuel*, Paris, Bibliothèque nationale de France, 2012. Disponible sur http://www.bnf.fr/documents/frsad_rapport_final.pdf

¹⁴³ BOURDON Françoise, ILLIEN Gildas, « A la recherche du temps perdu, retour vers le futur : CBU 2.0 », in 80e Congrès de l'IFLA, Lyon, mai 2014. Disponible sur <http://library.ifla.org/956/>

en plus autonomes, notamment parce qu'ils véhiculent des informations à forte valeur ajoutée qui ont été collectées et validées par des experts et qui ont un intérêt bien au-delà de l'environnement du catalogue. Les autorités sont par ailleurs particulièrement adaptées au fonctionnement des moteurs de recherche consistant à présenter à l'utilisateur des « entités » entretenant des relations les unes avec les autres.

Retenons qu'une tendance forte de la description bibliographique va donc consister à produire moins de données bibliographiques et davantage de données d'autorité. Les entités « œuvre » et « expression » sont en outre représentées par des autorités, ce qui a pour effet de provoquer un accroissement des notices d'autorités au détriment des notices bibliographiques.

Ce modèle conceptuel nécessite des outils adaptés, ce qui a conduit le monde anglo-saxon à élaborer un nouveau code de catalogage, RDA.

Un nouveau code de catalogage : RDA

Qu'est ce que RDA ?

RDA (*Resources Description and Access*) est un code de catalogage centré sur la notion d'information requise à la description d'une ressource, indépendamment de son format d'encodage. En effet RDA ne traite pas la question de la représentation de l'information. Par conséquent il peut être exprimé dans différents formats, par exemple dans un format MARC, ce qui est actuellement le cas pour les bibliothèques qui l'utilisent, ou en RDF, qui est la cible visée à terme afin de produire directement des données bibliographiques compatibles avec les standards du web sémantique.

Ce code de catalogage publié en 2010 est issu de la volonté de faire évoluer les règles de catalogage anglo-américaines, les AACR (*Anglo-American Cataloguing Rules*). Le processus de révision commence en 1997 lors de la Conférence internationale de Toronto sur l'évolution des AACR, afin de prendre en compte l'évolution des ressources et du contexte de l'information bibliographique et afin d'adapter les catalogues aux caractéristiques du web. Il est pris en charge par le JSC (*Joint Steering Committee for Development of RDA*) qui regroupe les bibliothèques nationales et agences bibliographiques des États-Unis, de l'Australie, du Canada, de la Grande-Bretagne et de l'Allemagne. Il s'agit donc à l'origine d'un contexte très anglo-saxon et d'un processus qui s'inscrit en dehors des canaux de normalisation internationale de l'IFLA. Par ailleurs RDA est mis à disposition sous la forme d'un site web payant (*RDA toolkit*) qui est régulièrement mis à jour.

Ce nouveau code de catalogage s'appuie sur FRBR et permet la description des ressources selon ce modèle. Il se présente sous la forme de 10 sections réparties en 2 grands groupes qui décrivent les attributs des entités (sections 1 à 4) et les relations entre elles (sections 5 à 10)¹⁴⁴.

RDA, par sa conception même, introduit un changement majeur dans l'appréhension du catalogage, que l'on peut qualifier avec Françoise Leresche de changement de paradigme¹⁴⁵. Les règles ISBD actuelles (et c'est le cas des AACR également) s'appuient en effet sur le modèle des fiches de catalogues imprimés, modèle

¹⁴⁴ JSC, *Resource description & access RDA / developed in a collaborative process led by the Joint Steering Committee for Development of RDA (JSC)*, Chicago, American Library Association, 2010

¹⁴⁵ LERESCHE Françoise, « Qu'est-ce que le RDA ? », [enregistrement vidéo], in 5 à 7 de l'ADBS, *RDA : le nouveau paradigme du catalogage ?*, Paris, 29 mars 2011. Disponible sur <http://www.adbs.fr/rda-le-nouveau-paradigme-du-catalogage-2-qu-est-ce-que-le-rda--102462.htm>

qui a été transposé quasiment à l'identique au moment de l'informatisation des catalogues. Cela explique l'existence de règles telles que la restriction du nombre de caractères (usage d'abréviations) ou la limitation de la transcription des noms d'auteurs à trois au maximum.

RDA permet de sortir de ce modèle contraint du fichier pour aller vers le modèle du web, à savoir un réseau comportant des liens, des relations entre notices, mais aussi à un niveau plus élémentaire entre données¹⁴⁶. Ainsi les vedettes sont remplacées par des points d'accès. De même, on abolit les limites liées au format contraint de la fiche telle que la restriction des informations de description.

Le grand avantage de ce code de catalogage est de permettre l'expression des données bibliographiques sur le web de données. Pour ce faire, l'utilisation des standards du web sémantique est requise : expression des données en RDF, utilisation d'URI pour l'identification pérenne. Un espace de nom recensant les éléments et référentiels de RDA a par conséquent été défini dans l'*Open Metadata Registry* (registre de métadonnées contenant des vocabulaires du web sémantique)¹⁴⁷.

Ainsi RDA est un code de catalogage résolument moderne, structuré à partir du modèle FRBR, adapté à l'exposition sur le web de données. Or la question de son adoption fait l'objet de grands débats dans la profession en Europe.

Adopter RDA ?

Précisons tout d'abord que l'adoption de RDA peut se faire selon trois scénarios. Dans le scénario 1 le catalogue s'appuie sur une structure de base de données relationnelle ou orientée-objet qui reflète parfaitement le modèle FRBR, c'est le scénario le plus abouti. Dans les scénarios 2 et 3, la structure des bases de données bibliographiques actuelles contenant des notices bibliographiques d'une part et des notices d'autorité d'autre part est maintenue. Ces deux types de notices sont liées entre elles dans le scénario 2 et ne le sont pas dans le scénario 3.

En France la majorité des bibliothèques ont déjà un modèle qui s'approche du scénario 2, car les notices d'autorité et les notices bibliographiques sont liées entre elles, grâce aux formats MARC utilisés (UNIMARC, INTERMARC) qui comportent des zones de liens. Cela n'est pas le cas aux États-Unis, le format MARC21 n'en disposant pas. De ce fait RDA est une évolution beaucoup plus fondamentale outre-Atlantique qu'en France, où seule l'adoption du scénario 1 aurait un intérêt.

Mais le groupe technique sur l'adoption de RDA en France a pointé plusieurs difficultés¹⁴⁸. Son rapport indique que RDA présente finalement une grande continuité avec les AACR et reflète par conséquent « une vision anglo-américaine du traitement de l'information » ce qui limite son ambition d'être adopté comme code international de catalogage.

Par ailleurs RDA repose sur une interprétation particulière du modèle FRBR, une « torsion » selon le rapport, probablement dans le but de faciliter la continuité avec les AACR. Ainsi la compréhension de la notion d'expression est hybride car

¹⁴⁶ *ibidem*

¹⁴⁷ <http://www.rdaregistry.info/>

¹⁴⁸ Groupe technique sur l'adoption de RDA en France, *Rapport du groupe technique sur l'adoption de RDA en France*, Rapport au Groupe stratégique pour sa réunion du 25 mars 2011. Disponible sur http://rda-en-france.enssib.fr/sites/rda-en-france.enssib.fr/files/Rapport%20au%20groupe%20strat%C3%A9gique%20pour%20r%C3%A9union%2025%20mars_v3_20110323_0.pdf

elle intègre des éléments qui relèvent de la manifestation. Par ailleurs RDA ne permet pas de décrire correctement certaines catégories de ressources telles que les périodiques, les ressources cartographiques ou encore l'image fixe.

L'adoption en France de RDA en l'état constituerait donc une régression (perte de précision dans le catalogage, analyses divergentes avec la culture française) mais aurait également un coût important dû à la nécessité de définir un profil d'implémentation (il faut opérer un choix entre de multiples options), rédiger des guides d'utilisation et former les catalogueurs. Néanmoins ne pas l'adopter consisterait à s'isoler sur le plan international, ce qui n'est pas souhaitable.

En conséquence, la position française consiste à affirmer la volonté d'adopter RDA à terme, sans préciser d'échéance, tout en maintenant une analyse française et en assurant la promotion¹⁴⁹.

Ce travail de tentative d'amélioration des règles de RDA est mené au niveau européen au sein d'EURIG, le Groupe d'intérêt européen sur RDA (*EUropean RDA Interest Group*) qui mesure l'adaptabilité de RDA dans les pays européens, dont les pratiques reposent sur les ISBD. EURIG émet des propositions d'évolution au JSC qui accepte ou non de les prendre en compte.

On peut légitimement se demander si EURIG a de réelles chances de faire aboutir ses positions, dans la mesure où ses propositions d'évolution font bien souvent l'objet d'un refus. Mais il faut noter que les travaux en cours sur la consolidation des modèles FRBR, FRAD et FRSAD pourraient aller dans le sens de la vision européenne, ce qui lui permettrait d'être mieux entendue.

Il faut cependant noter que la France ne se maintient pas totalement en dehors de RDA, des éléments du code sont adoptés et font l'objet de travaux d'implémentation dans les systèmes de catalogage et de campagnes de formation. Les modalités de cette transition seront détaillées dans le troisième chapitre.

Par ailleurs, comme l'indique Françoise Leresche dans un article intitulé « Avons-nous besoin de RDA pour mettre les catalogues de bibliothèques sur le web de données ? » l'adoption de RDA n'est pas une fin en soi. Le but recherché est par contre bien la « FRBRisation » des catalogues et leur exposition sur le web de données¹⁵⁰. Nous verrons dans la suite de ce travail comment ces priorités sont mises en œuvre.

Vers une révision des formats MARC

Le dernier point de la transition bibliographique concerne la révision des formats MARC. Ce chantier n'est pas encore complètement d'actualité en France, mais la réflexion est déjà en cours.

Les formats MARC, construits sur les ISBD, sont des formats datés, très peu répandus en dehors des bibliothèques, inadaptés au modèle FRBR et à la diffusion sur le web de données. D'autre part ils contribuent à enfermer les bibliothèques dans un écosystème informatique très spécifique qui les contraint à ne pouvoir faire appel qu'à un nombre restreint de fournisseurs.

¹⁴⁹ BnF. Pour les professionnels. RDA (Ressources: Description et Accès) [en ligne]. Paris, Bibliothèque nationale de France. Disponible sur http://www.bnf.fr/fr/professionnels/rda/s.rda_en_france.html?first_Art=non

¹⁵⁰ « FRBRiser les catalogues, les mettre sur le Web de données et passer de la récupération de notices à des liens entre données exposées sur le Web, telles sont aujourd'hui les priorités pour les bibliothèques françaises. » LERESCHE Françoise, *Avons-nous besoin de RDA pour mettre les catalogues de bibliothèques sur le web de données ? : la réflexion française sur l'évolution des catalogues et des règles de catalogage*, 2012. Disponible sur <http://www.bn.gov.ar/descargas/catalogadores/encuentro2011/ponencia-23-A-Leresche.pdf>

Notons cependant que les formats MARC sont parfaitement adaptés à l'usage qui est fait d'eux actuellement, à savoir la description d'un document. Ils permettent de rendre compte de manière fine des différents éléments d'information, de mettre en place des contrôles assez puissants et ils comportent des zones de liens. Cette finesse fait qu'il est difficile de rendre compte en RDF d'une information encodée en MARC.

Même dans l'hypothèse de l'adoption de RDA, il reste encore possible de travailler avec MARC, du moins dans une phase de transition. Les données bibliographiques peuvent être exposées sur le web de données a posteriori. Mais à terme se pose la question de la possibilité de produire des données bibliographiques directement en RDF. Or RDF est un formalisme présentant un haut niveau d'abstraction, et la complexité des données de bibliothèque est telle qu'il est nécessaire de construire des applications, des modèles pour les exprimer en RDF.

C'est ce qui occupe actuellement la Bibliothèque du Congrès qui a lancé en 2011 l'initiative BIBFRAME pour *Bibliographic Framework Initiative*, qui vise à définir un nouveau format pour les données des catalogues, en remplacement des formats MARC et en s'appuyant sur un modèle RDF. Ce nouveau modèle a pour ambition de rendre compte de tous les éléments de RDA tout en étant « général, assez léger pour être compris par les communautés du Web et applicable indépendamment de RDA ». ¹⁵¹ Cet objectif semble intéressant mais de premières critiques se font entendre sur la trop grande simplicité du modèle, sa non conformité au modèle FRBR et sa proximité trop grande avec MARC21 ¹⁵².

Signalons qu'OCLC propose une alternative à BIBFRAME nommée BiblioGraph et qui s'appuie sur schema.org, mais qui est encore trop récente pour que l'on dispose d'éléments d'évaluation ¹⁵³.

Mentionnons enfin le modèle FRBRoo, qui est issu d'un rapprochement entre le modèle FRBR et le modèle CIDOC CRM. Ce modèle, contrairement à BIBFRAME, répondrait à des besoins d'expression fine du MARC dans un format RDF.

Le modèle CIDOC ¹⁵⁴CRM (*Conceptual Reference Model*) est le modèle conceptuel de référence pour l'information muséographique. Il a été publié pour la première fois en 1998 et est devenu une norme ISO en 2006 (ISO 21127). « Il s'agit d'un modèle basé sur une approche orientée objet, c'est-à-dire sur la formalisation des relations entre des entités et des attributs de ces entités. On qualifie également le CRM CIDOC de modèle orienté événement, car il fait un usage intensif de la notion d'événement pour relier entre eux des objets, des personnes, des lieux, des activités, et d'autres notions conceptuelles. Même s'il a émergé à l'origine dans la sphère des musées, ce modèle se veut suffisamment abstrait pour pouvoir traiter de tout type de données culturelles ¹⁵⁵. »

¹⁵¹ LERESCHE Françoise, « Réunion annuelle d'EURIG », in Bulletin des bibliothèques de France, n° 6, novembre 2013. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2013-06-0077-004> (consulté le 10 octobre 2014)

¹⁵² <http://rda.abes.fr/2013/02/05/whats-wrong-with-bibframe/>

¹⁵³ <http://bibliograph.net/>

¹⁵⁴ CIDOC : Comité international pour la documentation du Conseil international des musées (ICOM-CIDOC)

¹⁵⁵ BERMES Emmanuelle, avec la collaboration d'Antoine ISAAC et Gautier POUPEAU, *Le Web sémantique en bibliothèque*, Paris, Ed. du Cercle de la Librairie, 2013

Un rapprochement entre le modèle FRBR des bibliothèques et le modèle CIDOC CRM des musées s'est opéré à partir de 2003, qui a abouti à la reformulation du modèle FRBR dans un formalisme orienté objet pour devenir FRBRoo.

Ce modèle fait aujourd'hui l'objet d'expérimentations au CIEPS (Centre international de l'ISSN) et à la BnF. Celle-ci a notamment entrepris de travailler à partir de FRBRoo sur un modèle permettant de rendre compte précisément de la finesse de description bibliographique contenue dans les formats MARC. Ce modèle, Figralida¹⁵⁶ est actuellement utilisé dans le cadre d'un programme de recherche sur la musique, Doremus¹⁵⁷.

On le voit, l'abandon des formats MARC n'est pas imminent. En attendant de pouvoir évaluer les propositions émanant d'outre-Atlantique et de construire de nouveaux modèles, ils vont encore nous accompagner pendant quelques années.

Ce tempo un peu lent, qui concerne aussi bien la révision de MARC que l'adoption du modèle FRBR ou du code de catalogage RDA, laisse le temps de préparer la transition sur le plan de l'organisation, car les modes de travail vont évoluer.

VERS DE NOUVEAUX MODES DE TRAVAIL ?

En effet des mutations sont en cours à la fois sur le plan de l'évolution des métiers des bibliothèques, en particulier l'activité de catalogage, qui touchent donc à l'identité professionnelle, mais aussi sur le plan de l'organisation du travail, par la façon dont celui-ci se distribue dans un réseau d'acteurs qui excède les frontières des bibliothèques.

Du catalogage au « cataliage »

Un rapport de l'Inspection générale des bibliothèques de 2012¹⁵⁸ évoque les mutations des métiers liées à l'environnement numérique des bibliothèques et les bouleversements occasionnés à la fois sur le plan des catalogues et des services. Parmi les grandes tendances soulignées par ce rapport on peut noter la mutualisation du signalement par les réseaux des agences bibliographiques, son corollaire la diminution des tâches de catalogage dans les établissements et enfin le développement de nouvelles compétences en particulier autour des métadonnées.

Ces bouleversements sont liés à l'évolution de la place du catalogue, qui de central devient une application parmi d'autres dans la bibliothèque, et tend de plus en plus à être un outil de back-office. Depuis plusieurs années de nombreuses interrogations se font entendre dans la profession sur l'avenir du catalogage, dont on voit une remise en cause avec l'activité de dérivation de notices et qui se poursuit avec le développement du web de données. Les catalogueurs sont-ils voués à disparaître ? Cela n'est pas souhaitable, nous allons voir pourquoi, mais l'activité de catalogage à moyen terme devrait avoir tendance à se concentrer sur certains acteurs. Et des changements dans le cœur de métier se dessinent d'ores et déjà.

¹⁵⁶ Cet acronyme signifie Finesse de la GRANularité des LIaisons entre Données librement Accessibles ou FIne-GRained Linked (open) Data. Le modèle FiGraLiDa a été développé par Patrick Le Bœuf.

¹⁵⁷ http://www.bnf.fr/fr/la_bnf/prog_anr.html

¹⁵⁸ AROT Dominique, CARBONE Pierre, CLAUD Joëlle, MICOL Charles, LECOQ Benoît, *Quels emplois dans les bibliothèques ? État des lieux et perspectives*, Inspection générale des bibliothèques, rapport n°2012-020, mars 2013. Disponible sur http://cache.media.enseignementsup-recherche.gouv.fr/file/2013/20/3/Quels_emplois_dans_les_bibliotheques_Etat_des_lieux_et_perspectives_247203.pdf

Ainsi le travail de catalogage consiste de plus en plus non pas à créer une information originale mais à rechercher et adapter une information existante. Cette tendance se manifeste par exemple dans la mise à jour en 2012 de la fiche de poste de Gestionnaire de collection de la BnF qui fait apparaître un nouvel item dans la liste des activités du catalogueur : « récupérer des métadonnées documentaires externes, les contrôler et, le cas échéant, les compléter¹⁵⁹ ». Le métier de catalogueur s'achemine en effet vers celui de gestionnaire de métadonnées. Cette mutation nécessite l'acquisition de nouvelles compétences : « Les activités de récupération, de versement ou de « liage » tous azimuts vont requérir des compétences et une ingénierie particulières qui constituent un axe fort d'évolution du métier - probablement le plus important à court et moyen terme¹⁶⁰. » Ce glissement du catalogage vers la création de liens vers des ressources existantes a conduit les anglo-saxons à forger le terme de « *catalinking* », traduit en français par « cataliage ». Gildas Illien souligne également l'importance de la dimension de négociation et d'administration de ces processus dans le futur métier de catalogueur. Autre facette de l'évolution du métier : « le traitement des documents portera moins sur leur description matérielle que sur l'analyse de leur contenu, en lien avec la maintenance de vocabulaires et de référentiels d'autorité partagés¹⁶¹ ».

L'activité du catalogueur va également évoluer grâce à la place prépondérante que vont prendre les algorithmes, sur laquelle il peut être utile de s'attarder un instant. En effet, autant du côté de l'ABES que de la BnF se développent des outils sophistiqués visant à aider le catalogueur. Cette nouvelle approche peut être déstabilisante au regard de la forte culture d'indexation manuelle qui domine dans les bibliothèques. Elle peut aussi générer la crainte que le travail manuel soit supplanté par les outils. Mais il s'agit plutôt d'instaurer une forme de complémentarité entre les deux.

Nous n'avons pas encore évoqué le projet data.bnf.fr, l'application d'exposition sur le web de données des données bibliographiques de la BnF dont il sera question au cours du troisième chapitre, mais anticipons sur un point. L'une des fonctions de data.bnf.fr consiste à apporter une couche de FRBRisation au catalogue général de la BnF en créant des liens entre une œuvre et ses manifestations, ce qui passe par le développement d'algorithmes de rapprochement. Ces algorithmes sont ensuite testés par des catalogueurs qui seuls ont l'expertise métier pour évaluer la qualité des liens créés automatiquement. Ceci constitue un autre axe d'évolution du métier. Comme le souligne Benjamin Duhamel dans son rapport de stage effectué au sein du service en charge de data.bnf.fr, « il s'agit de développer une compétence d'ingénierie des données portant sur les algorithmes eux-mêmes dont les experts des données améliorent les spécifications en l'enrichissant de règles métiers¹⁶² ».

On retrouve aussi cette interaction entre le catalogueur et l'algorithme dans le projet du hub de métadonnées de l'ABES. Là encore il ne s'agit pas de remplacer l'homme par la machine. On pourrait être tenté en effet d'imaginer pouvoir se reposer sur ce genre d'outils pour gérer l'énorme masse de données à laquelle sont confrontées aujourd'hui les bibliothèques. Mais cet outil à forte valeur ajoutée comporte aussi des limites. Tout d'abord l'approche du hub de métadonnées a un

¹⁵⁹ BnF, *Référentiel des emplois et des compétences*, 2012. Disponible sur http://www.bnf.fr/emploi/pdf/referentiel_emplois_compétences.pdf

¹⁶⁰ ILLIEN Gildas, « Décrire les objets du savoir, les nouveaux paradigmes du catalogage », *op.cit.*

¹⁶¹ *ibidem*

¹⁶² DUHAMEL Benjamin, *Les technologies du web sémantique et du record linkage au service de data.bnf.fr et du Linked Open Data culturel : Étude sur les nouveaux paradigmes informationnels*, mémoire de Master à l'Université Lille 3 sous la direction de Gildas Illien et Gérard Kembellec, 2014.

coût, d'où la nécessité d'y recourir lorsque cela est pertinent par rapport à une intervention humaine. Ainsi l'ABES a choisi de donner la priorité aux corpus des licences nationales qui comportent de gros volumes de données et dont le traitement est plus facilement automatisable que des petits corpus hétérogènes. Par ailleurs il s'agit avant tout de fournir une aide au catalogage, en allant le plus loin possible dans les traitements automatiques, avant de solliciter une intervention humaine. Ainsi dans le cadre de l'exploitation des *e-books*, le hub réalise des enrichissements (lien avec les autorités IdRef ou Rameau, détection automatique des langues des résumés). Puis, pour la suite des traitements qui ne peuvent faire l'économie d'une intervention humaine, une zone du Sudoc est alimentée pour indiquer aux catalogueurs quelles améliorations restent à apporter, comme une sorte de post-it comportant une liste des tâches à réaliser. Il s'agit là d'une première tentative d'interaction entre le hub et le réseau des catalogueurs. L'équipe du hub réfléchit à la manière dont à l'avenir le réseau pourrait collaborer au hub plus directement¹⁶³.

Un autre projet de l'ABES va dans cette même direction consistant à proposer des outils qui pourraient appuyer le catalogueur dans son activité. Il s'agit de Qualinca¹⁶⁴ pour « qualité et interopérabilité de grands catalogues documentaires », un projet de recherche fondamentale financé par l'Agence nationale de la recherche qui prend la suite du projet SudocAd¹⁶⁵. Il associe l'ABES, l'INA (Institut National de l'Audiovisuel) et trois laboratoires¹⁶⁶. L'objectif de Qualinca est de qualifier et améliorer le niveau de qualité d'une base documentaire existante. Il s'agit de mesurer la qualité des liens et de générer de nouveaux liens vers des autorités. Le projet s'appuie sur une approche théorique consistant à exprimer des règles logiques qui permettront par exemple de déterminer quelle est la bonne autorité¹⁶⁷. Le but est à terme de réaliser des prototypes de liage automatique et d'aide à la décision pour le catalogueur. Ceux-ci seront évalués « en comparant décisions automatiques et décisions humaines. Le moment venu, l'ABES sollicitera la participation des catalogueurs du réseau Sudoc à ce protocole d'évaluation¹⁶⁸. »

Dans tous les cas évoqués ci-dessus, une question fondamentale est celle du niveau de qualité de ce qui est produit par les algorithmes. Il doit dans la mesure du possible être comparable à ce que produisent les catalogueurs, de manière à ne pas générer deux niveaux de qualité incohérents. Cela nécessite de définir un seuil de confiance acceptable, mais aussi de vérifier constamment ses variations dans le temps, afin de maintenir un rapport satisfaisant entre gains et risques encourus.

Une nouvelle forme d'interaction homme machine se dessine donc avec l'émergence d'outils sophistiqués en appui au travail de catalogage, qui reste indispensable et se recentre sur des opérations à forte valeur ajoutée.

Concluons sur l'évolution du métier par deux remarques.

Tout d'abord, la réussite du web de données repose sur une grande qualité des données, et la réputation de fiabilité des bibliothèques en est aussi un des enjeux. Ce point peut constituer un levier de valorisation pour la fonction de catalogage à l'heure où

¹⁶³ NICOLAS Yann, « Focus sur le hub de métadonnées », *op.cit.*

¹⁶⁴ <http://www.lirmm.fr/qualinca/>

¹⁶⁵ <http://www.abes.fr/Sudoc/Focus-sur-les-etudes-et-projets/SudocAD>

¹⁶⁶ Le Laboratoire d'informatique, de robotique et de microélectronique de Montpellier (LIRMM), le Laboratoire d'informatique de Grenoble (LIG) et le Laboratoire de recherche en informatique d'Orsay (LRI)

¹⁶⁷ Exemple de règle logique : Il est impossible qu'une personne publie quelque chose avant d'être née. Ou encore il est peu probable qu'une même personne soit à la fois poète et mathématicien (exemples proposés par Michael Jeulin et Yann Nicolas).

¹⁶⁸ <http://www.abes.fr/Projets-en-cours/Qualinca>

celle-ci apparaît comme en déclin. Cette production de données va concerner de plus en plus les données d'autorité, travail à forte valeur ajoutée demandant un réel investissement de la part du catalogueur. « Alors que les notices bibliographiques se limitent souvent à reproduire et organiser des informations objectives présentes sur les documents décrits, les notices d'autorité sont le fruit d'un travail original d'identification de la part du catalogueur qui recherche puis confronte des sources diverses, et engage sa responsabilité sur les choix opérés.¹⁶⁹ »

De plus, il est capital que les bibliothèques conservent la maîtrise de leurs données. Nous avons vu que la tendance des nouveaux SGB consistait à offrir des services « tout compris », incluant notamment des données bibliographiques et d'autorité. Or il peut être dangereux de laisser aux fournisseurs de SGB la maîtrise des données des bibliothèques. C'est ce que rappelle l'ABES avec sa stratégie d'indépendance des données qui milite « pour des outils vides »¹⁷⁰.

Une nouvelle répartition du travail entre les acteurs

Une autre tendance que l'on voit se dessiner et qui vient compléter le renforcement du travail en réseau est une répartition du travail entre différents acteurs, qui s'opère à plusieurs niveaux. En effet il est clair que les bibliothèques ne peuvent plus se partager le signalement des ressources comme au temps du contrôle bibliographique universel. Le CBU reposait sur le modèle de l'imprimé, une publication étant alors généralement identifiable à un pays. Or avec le numérique, outre l'explosion du nombre de ressources à traiter, les frontières entre pays sont rendues extrêmement poreuses. Comment assigner des responsabilités nationales sur des contenus créés de manière collaborative à l'échelle internationale ? Le signalement doit donc passer par de nouveaux modes de répartition des activités entre les bibliothèques, à l'échelle nationale, internationale, mais aussi avec d'autres acteurs.

A l'échelle nationale il apparaît que les deux agences bibliographiques sont vouées à tenir un rôle particulier vis-à-vis de l'ensemble des bibliothèques du territoire. Cette idée revient à la fois dans le rapport de l'IGB mentionné plus haut¹⁷¹ et dans celui de l'ABES et la BnF au CSB qui prône que les agences s'affirment comme « les opérateurs de réservoirs de métadonnées de référence dans lesquels chaque établissement viendra puiser au lieu d'en dupliquer la production localement¹⁷² ». Cela permettra de provoquer un allègement et une transformation du travail des établissements, qui pourront « se concentrer sur le traitement des données locales, la description de fonds spécialisés et l'enrichissement des données par la création de liens¹⁷³ ». Dans le cas de l'enseignement supérieur, ce qui se dessine avec le hub de métadonnées est une nouvelle forme de coproduction des données réalisée conjointement par le réseau et l'ABES.

¹⁶⁹ BOURDON Françoise, ILLIEN Gildas, ROCHE Mélanie, « Le catalogue des temps modernes, entre discipline et dissémination », in BARBIER Frédéric, DUBOIS Thierry, SORDET Yann, *De l'argile au nuage : une archéologie des catalogues, catalogue d'exposition*, Paris, Bibliothèque Mazarine (13 mars 2015 - 13 mai 2015), Paris, Bibliothèque Mazarine, Bibliothèque de Genève, Éditions des Cendres, 2015.

¹⁷⁰ NICOLAS Yann, « Focus sur le hub de métadonnées », *op.cit.*

¹⁷¹ « On peut se demander si l'essentiel du catalogue original (et de l'indexation) ne va pas à l'avenir reposer sur quelques établissements agissant comme agences prestataires pour le réseau des bibliothèques (BnF, ABES, CADIST) ».

¹⁷² ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique*, *op.cit.*

¹⁷³ *ibidem*

A l'échelle internationale on peut d'une certaine manière retrouver un fonctionnement similaire au CBU, mais de manière moins formalisée. Cela passe par la dérivation de notices, l'utilisation et l'alimentation des catalogues collaboratifs (Worldcat) et des référentiels d'autorités tels que VIAF¹⁷⁴ (*Virtual International Authority File*) ou ISNI¹⁷⁵ (*International Standard Name Identifier*). Dans le champ de l'enseignement supérieur et de la recherche, ce que fait l'ABES avec Bacon en diffusant sur le web de données les métadonnées des corpus d'éditeurs francophones contribue à cette « division internationale du travail¹⁷⁶ ».

La grande évolution par rapport au CBU consiste à utiliser les données produites par d'autres communautés professionnelles. Cela a été évoqué pour les données produites par les éditeurs, à la fois pour la BnF dans le cadre du dépôt légal, et pour l'ABES dans le cadre du hub de métadonnées, concernant certains corpus des licences nationales. Ce point illustre d'ailleurs bien la complémentarité des agences au vu de leurs missions respectives.

Cette tendance a vocation à se réaliser aussi avec les autres acteurs du monde de la culture (interprofession du livre, musées, archives). C'est dans cette perspective qu'est développé le modèle FRBRoo. L'alignement des référentiels propres à différents domaines culturels est également un moyen de lier ses données à d'autres acteurs. C'est le cas avec le Thésaurus W des Archives nationales de France qui est lié avec Rameau¹⁷⁷. Les bibliothèques ont tout à y gagner, comme le souligne le rapport du W3C, elles « récolteront les fruits de l'ouverture de leurs données, à savoir des contributions de valeur provenant d'autres communautés. Ajouter des liens vers des données venant de biographes ou de généalogistes, par exemple, pourrait enrichir la description des ressources documentaires avec des données généralement non fournies par les bibliothèques, et pourrait grandement améliorer la découverte et la navigation des collections des bibliothèques¹⁷⁸. » Le principe qui sous-tend l'utilisation du web de données est souvent exprimé de la façon suivante : « penser global, agir local ». Cela signifie qu'il est possible de procéder à une description fine et précise dans son domaine de spécialité, tout en préservant la possibilité d'être présent et visible de toute la communauté par le biais des alignements.

Le web de données permet en outre d'envisager cet élargissement du périmètre des acteurs qui coopèrent de manière exponentielle, comme le montre le nuage du LOD précédemment évoqué, le premier gisement de données auquel se lient les bibliothèques étant Wikipédia.

Dans leur contribution au congrès de l'IFLA sous-titrée « CBU 2.0 », Françoise Bourdon et Gildas Illien tout en constatant que les bibliothèques ne sont plus en mesure de décrire toute l'information à elles seules, estiment que l'idéal du CBU, « faire connaître à chacun ce qui a été produit partout dans le monde moyennant un partage international des tâches¹⁷⁹», reste valable. Mais cela nécessite de s'appuyer sur cette nouvelle répartition des tâches avec des acteurs excédant de loin le monde des bibliothèques, y compris les usagers du web. Ce CBU nouvelle version repose sur certains principes, l'un d'entre eux étant la nécessité de maintenir l'interopérabilité entre bibliothèques, ce qui passe par la définition de formats d'échanges tels que BIBFRAME ou schema.org. Il est également primordial d'adopter les standards du web sémantique et

¹⁷⁴ <http://viaf.org/>

¹⁷⁵ <http://www.isni.org/>

¹⁷⁶ NICOLAS Yann, « Focus sur le hub de métadonnées », *op.cit.*

¹⁷⁷ BERMÈS Emmanuelle, « Bibliothèques, archives et musées : l'enjeu de la convergence des données du patrimoine culturel », in *Documentaliste-Sciences de l'information*, vol.48, n°4, décembre 2011, p.45-47

¹⁷⁸ GROUPE D'INCUBATION W3C, *op.cit.*

¹⁷⁹ BOURDON Françoise, ILLIEN Gildas, *op.cit.*

de participer à sa gouvernance, seul moyen de faire sortir les bibliothèques de leur univers cloisonné.

Cette nouvelle configuration consistant à lier ses données à celles d'autres acteurs permet de répondre aux enjeux actuels liés au signalement, mais elle pose aussi un certain nombre de questions qu'il faut avoir à l'esprit.

L'idée qu'il sera nécessaire de faire moins de catalogage, dans un nombre restreint d'institutions, peut inciter à réduire les ressources dédiées à ces fonctions dans les établissements, ou encore à réduire les moyens consacrés à la formation au catalogage. Mais cette tendance doit être suivie avec modération. Si toutes les bibliothèques se mettaient à se reposer sur les autres pour produire des données, le système ne tiendrait plus. Il faut donc veiller à maintenir et coordonner une activité de catalogage.

D'autre part, le fait de ne plus contrôler la production de données mais de se lier avec des jeux de données existants comporte des risques. Lors de l'atelier du groupe d'intérêt sur le web sémantique qui s'est tenu lors du dernier congrès de l'IFLA, a été évoquée la question de la confiance dans les jeux de données disponibles sur le web, ceux-ci pouvant s'avérer instables dans la durée. Il faut donc choisir avec discernement les jeux de données auxquels on se lie. Quelques critères ont été évoqués pour définir le degré de confiance d'un jeu de données : « leur origine institutionnelle, [...] l'utilisation d'identifiants pérennes et standardisés, la qualité et la fiabilité des données sources, le marquage de l'information de provenance, la documentation et la mise à jour régulière des jeux de données¹⁸⁰. »

Les changements que nous venons de décrire ne sont pas encore survenus. Il s'agit d'une transition qui va vraisemblablement s'étirer sur une dizaine d'années et qui reste à construire. Voyons selon quelles modalités l'ABES et la BnF s'y préparent.

¹⁸⁰ BnF, *L'IBN à l'IFLA : Rapport de mission. Congrès de Lyon et Conférences satellites*, 2014.

OUVRIR LES DONNÉES : LES STRATÉGIES MISES EN ŒUVRE

Les chapitres précédents ont montré que la communauté des bibliothèques se trouve actuellement au cœur de deux mouvements convergents, celui de la transition bibliographique et celui de l'ouverture des données publiques. L'objectif de ce chapitre est de décrire les moyens mis en œuvre par les deux agences pour participer à cette mutation et en particulier la façon dont elles se concertent et coopèrent pour y parvenir.

Les modalités de cette concertation sont définies dans la convention de coopération qui les lie et qui a été renouvelée en juin 2014. Au-delà de l'échange de données évoqué plus haut, cette convention fixe que « l'ABES et la BnF entendent poursuivre ou initier d'autres collaborations et des expérimentations conjointes [...]. Le développement de ces collaborations s'inscrit dans une démarche prospective qui vise à la mise en œuvre d'une stratégie nationale de mutualisation, de diffusion, de valorisation et de réutilisation des métadonnées qu'elles produisent et qu'elles gèrent, au service de la communauté nationale et internationale. »

Le document pose les principes pour cette collaboration :

- « Assurer l'ouverture juridique et la gratuité de réutilisation des métadonnées », affirmant l'adhésion aux principes de *l'open data*.
- « Participer au développement d'un écosystème d'opérateurs publics et privés dans le domaine des métadonnées », s'inscrivant ainsi dans le sillage des préconisations du rapport Lescure.
- « Expérimenter et favoriser l'adoption par la communauté professionnelle des standards et technologies du web sémantique », affirmant par là leur volonté conjointe d'être à l'avant-garde du mouvement de dissémination des données bibliographiques sur le web.

Plusieurs axes de travail sont détaillés dans la convention. Certains ne seront pas abordés dans le cadre de cette étude, par exemple la collaboration sur un outil national de production au format EAD ou encore la dématérialisation des demandes ISSN de l'ABES au centre ISSN France. Nous retenons par conséquent les deux axes suivants :

- « Concertation sur les formats et règles de production », thème au cœur de la transition bibliographique qui comprend l'adoption du modèle FRBR, du code de catalogue RDA et les réflexions sur la révision des formats MARC.
- « Constitution ou modernisation de fichiers nationaux d'autorité partagés », soit l'un des axes stratégiques nécessaires à la réalisation du contrôle bibliographique universel.

Pour mettre en œuvre cette concertation, des agents de l'ABES et de la BnF, principalement du département de l'IBN, se réunissent deux fois par an pendant deux jours pour travailler sur l'ensemble des chantiers en cours.

Rappelons-le, la stratégie adoptée en France dans la transition bibliographique consiste à donner la priorité à la FRBRisation des catalogues et à leur exposition sur le web de données. L'étape consistant à réviser les modes de production des données, par exemple en produisant directement du RDF, fait l'objet de réflexions mais qui sont encore très largement prospectives. Chacune des agences s'est donc lancée dans l'exposition des données bibliographiques sur le web de données, selon des modalités un

peu différentes qui reflètent aussi la culture qui leur est propre. Elles travaillent en outre à cet objectif commun qu'est la constitution de fichiers d'autorités nationaux : personnes et collectivités, sujets et enfin œuvres, ce qui permet de poser la question : comment FRBRiser les catalogues ? Enfin, la question de l'adoption de RDA, qui s'annonce relativement longue compte tenu de l'inadaptation actuelle du code au contexte français, est instruite parallèlement. Elle est traitée dans une nouvelle structure, qui du fait de l'imbrication des sujets prend également en compte les autres aspects de la transition bibliographique. Ses attributions vont bien au-delà de la normalisation et permettent d'envisager la question de la conduite du changement.

S'INSÉRER DANS LE WEB DE DONNÉES

« L'ABES et la BnF considèrent que le Web de données est la destination à privilégier pour les métadonnées des bibliothèques », celui-ci constituant « un espace documentaire unifié, ouvert et global, capable de tisser des liens entre données à un niveau très fin de description et de désenclaver les données de leurs silos de production enfouis dans le Web invisible¹⁸¹ ». Autrement dit l'interopérabilité peut aujourd'hui se réaliser au niveau des données et non plus des systèmes d'information.

Les deux agences se sont lancées dans le web de données à peu d'années d'intervalle, en 2008 pour l'ABES avec Calames et 2011 pour la BnF avec le lancement de data.bnf.fr. Par ailleurs elles ont toutes deux placé leurs données sous la licence ouverte de l'État en 2012 et 2014, mettant ainsi en œuvre l'ouverture juridique. Cette section vise à expliquer un peu plus dans le détail comment est mise en œuvre l'ouverture technique et à comparer les stratégies choisies par les deux agences.

Avant cela posons quelques préalables techniques. RDF est un standard du web de données qui se décline dans plusieurs formats. RDF-NT (*N-Triples*) en est l'exemple le plus simple dans la mesure où il repose sur une syntaxe élémentaire développant chaque triplet. RDF-N3 ou Turtle (*Terse RDF Triple Language*) adopte une structure plus compacte, tandis que RDF/XML encapsule les informations dans des balises XML (*eXtensible Markup Language*), ce qui en fait un format plutôt destiné aux machines¹⁸². JSON LD (*JavaScript Object Notation for Linked Data*) est une alternative plus légère à RDF/XML. Enfin RDF/a est un format qui permet d'embarquer des données dans des pages HTML afin d'améliorer la présentation des résultats par les moteurs de recherche et qui s'appuie sur le même type de technologie que schema.org.

Il existe plusieurs techniques pour exposer des données en RDF :

- La première consiste à embarquer du RDF dans des pages HTML en y incluant des données en RDF/a. Il s'agit en général d'une stratégie complémentaire à celles qui vont suivre, son intérêt étant uniquement au niveau de la visualisation des résultats lors d'une recherche.
- On peut également mettre à disposition un lot de données à télécharger, que l'on appelle un dump. Un dump RDF constitue un export de la base de données à un instant T. Étant particulièrement long à générer les dumps ne

¹⁸¹ ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique*, op.cit.

¹⁸² <http://documentation.abes.fr/sudoc/manuels/pdf/sudoc-rdf.pdf>

sont pas mis à jour très fréquemment, ce qui pose des inconvénients en termes de fraîcheur des données. De plus le fait de disposer en un seul bloc de l'ensemble d'une base n'est pas très commode et ne peut servir qu'à des usages bien précis.

- Une autre possibilité consiste à mettre à disposition les données via un SPARQL endpoint, un service d'interrogation des données RDF. Cela permet d'interroger une base de manière dynamique en utilisant le langage de requête SPARQL, et fournit en retour une sélection des données. Par exemple : la liste de toutes les œuvres auxquelles a contribué telle personne. Si le requêteur SPARQL s'appuie sur des données issues d'un dump, le problème de la fraîcheur des données n'est pas résolu.
- Enfin, on peut mettre à disposition les pages HTML dans l'un des formats RDF énumérés plus haut (XML, JSON, N3, NT...) à la volée, donc en temps réel. Ce dispositif permet d'avoir des données parfaitement à jour, par contre il ne permet pas de récupérer de grandes quantités de données. En fonction des paramètres du navigateur, celui-ci affiche une page HTML (pour les humains), une page en RDF (pour les programmes) ou encore une page HTML composée de métadonnées en RDF/a (pour les robots des moteurs de recherche). Ce mécanisme s'appelle la négociation de contenu. Il consiste à prendre en compte les paramètres du navigateur pour faire varier le contenu affiché à partir d'une même URL, (par exemple pour une ressource en différentes langues). On peut trouver une illustration très éclairante de ce point sur le blog technique de l'ABES¹⁸³.

L'approche de l'ABES

Principes et éléments techniques

La première expérience de l'ABES avec le web de données a lieu en 2008 avec son application Calames. Il s'agit alors d'utiliser le RDF/a pour améliorer l'affichage des résultats des liens vers le catalogue dans les moteurs de recherche¹⁸⁴. En 2010 au moment du lancement d'IdRef, la base des autorités du Sudoc, les données sont rendues disponibles en RDF/XML. C'est ensuite au tour du Sudoc, en 2011. Ses données sont exposées en RDF/XML et ses pages embarquent des métadonnées conformes à schema.org. Le portail theses.fr associe également lors de sa création en 2011 la technique RDF/XML combinée à RDF/a. Les principales bases de l'ABES sont donc exposées sur le web de données, selon des modalités qui varient un peu de l'une à l'autre. Mais elles sont avant tout liées entre elles autour d'IdRef.

Les bases Sudoc et IdRef disposent d'un dump mais il n'est pas officiellement mis à la disposition du public. Un requêteur SPARQL est en train d'être mis au point, avec la particularité qu'il n'interrogerait pas le dump mais directement la base de données, ce qui comporte un avantage en termes de fraîcheur des données mais peut poser des problèmes de performance.

Les URI pérennes utilisées par l'ABES pour exposer ses données sont construites pour le Sudoc et IdRef sur la base des identifiants de notices¹⁸⁵ (PPN) et sur d'autres identifiants pour Calames et theses.fr.

¹⁸³ <http://pункtokomo.abes.fr/2011/07/04/le-sudoc-sur-le-web-de-donnees/>

¹⁸⁴ NICOLAS Yann, « Calames, et après ? », in *Bulletin des bibliothèques de France*, n° 6, novembre 2008. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2008-06-0029-005>

¹⁸⁵ <http://www.idref.fr/02669719X>

L'ABES applique avec conviction le principe de l'utilisation de vocabulaires existants : « pas de vocabulaire maison¹⁸⁶ ». Ces vocabulaires sont Dublin Core, FOAF, RDA, déjà cités, ou encore *Bibliographic ontology*¹⁸⁷ et ISBD¹⁸⁸. Ce choix s'applique y compris lorsque ces vocabulaires s'avèrent insuffisants. Par exemple, il n'existe pas de listes des codes de fonction UNIMARC exprimée en RDF. L'ABES a fait le choix d'exploiter la liste MARC21 et pour les quelques cas où les besoins ne sont pas couverts, préfère sacrifier la précision en utilisant un terme générique¹⁸⁹ tel que « dcterms:creator ».

Les données de l'ABES sont liées avec d'autres référentiels : la classification Dewey, MeSH (le thésaurus biomédical maintenue par la Bibliothèque nationale de médecine des Etats-Unis), Geonames ; et par l'intermédiaire de VIAF, dont le fonctionnement sera détaillé plus loin, à de nombreux référentiels notamment ceux de bibliothèques nationales.

Quelles réutilisations ?

L'exposition des données a un impact sur le référencement des notices par les moteurs de recherche, puisqu'elle leur permet de sortir du web profond. Par ailleurs, à travers les partenaires auxquels l'ABES s'est liée, ses données peuvent être trouvées par rebond, par des hommes ou des machines. Le nuage du LOD¹⁹⁰ fait bien apparaître les différents jeux de données de l'ABES au cœur d'un ensemble de liens vers ces autres jeux de données.

Les réutilisations par des développeurs ou pour d'autres applications sont parfois difficiles à mesurer. En effet le principe de l'ouverture des données repose sur le fait que les utilisateurs peuvent rester inconnus, puisque aucune inscription ou déclaration n'est requise. Pour les informer mais aussi tenter de susciter un retour de leur part, l'ABES met à leur disposition différents canaux de communication. La documentation concernant le Sudoc en RDF est disponible en ligne¹⁹¹. L'ABES publie régulièrement sur son blog technique *Punktokomo* des billets relatifs à l'exposition des données et ce blog est ouvert aux commentaires afin de constituer un espace de discussion.

Cependant les utilisateurs qui font partie du réseau Sudoc font en général connaître leurs réutilisations. On peut citer le cas de Philippe Bourdenet à l'Université du Maine qui a développé une fonctionnalité permettant, à partir d'une notice d'autorité du catalogue local, d'afficher des informations complémentaires sur l'auteur, à travers un web service de l'ABES qui renvoie les données exposées d'IdRef. Cette fonction « en savoir plus sur cet auteur » permet d'afficher la liste de toutes les contributions de celui-ci classées par codes de fonction (auteur, préfacier, éditeur, traducteur...)¹⁹². Deux autres réutilisations ont été présentées lors des journées ABES 2014. La première par Julien Sicot de l'Université Rennes 2, qui utilise les données exposées du Sudoc notamment pour réaliser un contrôle qualité des données¹⁹³. La seconde est un outil d'édition de bibliographie

¹⁸⁶ NICOLAS Yann, « L'Abes engage ses forces et ses données », in *Documentaliste-Sciences de l'information*, vol.48, n°4, décembre 2011, p.51

¹⁸⁷ <http://bibliontology.com/>

¹⁸⁸ <http://metadatareregistry.org/schema/show/id/25.html>

¹⁸⁹ Entretien avec Michaël Jeulin et Yann Nicolas, ABES, 15 octobre 2014

¹⁹⁰ <http://lod-cloud.net/>

¹⁹¹ <http://documentation.abes.fr/sudoc/manuels/pdf/sudoc-rdf.pdf>

¹⁹² BOURDENET Philippe, « Enrichir le catalogue local avec des données exposées : interactions avec le catalogue national via web services », in 80^e Congrès de l'IFLA, Session 86 - *Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?*, Lyon, 18 août 2014. Disponible sur <http://library.ifla.org/815/>

¹⁹³ <http://fr.slideshare.net/abesweb/jabes14-julien-sicotutiliserwebservicesabes-35302040>

(DoMyBiblio) développé par Yves Tomic de l'Université Paris Sud et qui s'appuie sur le web service de l'ABES « Sudoc en RDF¹⁹⁴ ».

Enfin, un autre utilisateur de ces données en RDF est l'ABES elle-même, pour ses outils internes. On a vu que le hub fonctionnait à partir d'une modélisation en RDF des données fournies par les éditeurs dans leur format natif. Le hub a en partie pour destination la mise à disposition des données qu'il produit sur le web de données. Mais pour le moment il est surtout utilisé pour traiter les données et les reverser dans le catalogue et dans la base de connaissances. Les données en RDF sont également utilisées pour des besoins internes dans le cadre du projet Qualinca.

Signalons enfin que la stratégie de l'ABES se décline aussi beaucoup autour de la fourniture de micro web services¹⁹⁵. Un web service est une API qui utilise les standards du web (le protocole HTTP). Un micro web service est un service simple, destiné à une seule tâche, par exemple localiser des documents à partir de leur identifiant (service Multiwhere¹⁹⁶). Une page dédiée aux développeurs donne toutes les informations nécessaires pour mettre en place ces web services¹⁹⁷. Il s'agit donc de faciliter la réutilisation des données en apportant des services autour de ces données. On peut considérer que cette approche est complémentaire de l'ouverture des données en tant que telle.

L'approche de la BnF

Principes et éléments techniques

L'application data.bnf.fr¹⁹⁸, conçue au départ dans une logique d'expérimentation, est lancée en juillet 2011. S'appuyant sur les standards du web sémantique elle se présente sous la forme d'un site web qui constitue une sorte de porte d'entrée à différents catalogues de la BnF : le catalogue général, celui des archives et manuscrits, et enfin Gallica. Ceux-ci forment en effet des silos qui ne communiquent pas entre eux, notamment parce qu'ils sont construits sur des formats différents (MARC, EAD).

Data.bnf.fr regroupe les ressources figurant dans ces catalogues dans des pages « auteur », « œuvre », « thème » ou « date ». Ainsi sur la page d'une œuvre toutes les versions qui en sont conservées à la BnF sont rassemblées (manuscrits, éditions électroniques, versions numérisées, adaptations audiovisuelles). Cette approche est donc tout à fait en cohérence avec le modèle FRBR. Ce projet, connu au début sous le nom de « pivot documentaire », permet de donner de la cohérence à l'offre documentaire de la BnF. Il illustre également comment le web de données permet d'atteindre l'interopérabilité sans toucher aux applicatifs.

Un autre objectif du projet est l'amélioration du référencement des ressources de la BnF sur le web, les pages de data.bnf.fr étant visibles des moteurs de recherche. « L'internaute n'est plus obligé de connaître, a priori, les différentes bases de recherche de la BnF, mais retrouve directement la ressource pertinente sur le web, en passant par les moteurs de recherche¹⁹⁹ ». Data.bnf.fr est donc un outil qui apporte de l'audience aux

¹⁹⁴ <http://punktokomo.abes.fr/2014/02/18/domybiblio/>

¹⁹⁵ <http://punktokomo.abes.fr/2011/07/04/les-micro-web-services-abes/>

¹⁹⁶ <http://documentation.abes.fr/sudoc/manuels/administration/aidewebservices/multiwhere.html>

¹⁹⁷ <http://www.abes.fr/Access-direct-a/Pour-les-developpeurs>

¹⁹⁸ <http://data.bnf.fr/>

¹⁹⁹ SIMON Agnès, WENZ Romain, « Des outils automatiques pour le signalement en bibliothèque : expérimentations autour du projet data.bnf.fr ». in *Bulletin des bibliothèques de France*, n°5, septembre 2012. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2012-05-0039-008>

différents catalogues de la BnF. Le taux de conversion de data.bnf.fr vers une autre application de la BnF est d'environ 75% en octobre 2014. De ce fait l'interface est assez sobre²⁰⁰, le but n'étant pas que l'internaute reste sur ces pages mais de le diriger vers les catalogues.

Les pages de data.bnf.fr sont générées automatiquement à partir des données des catalogues, dont la couverture s'effectue de manière progressive. En novembre 2014 plus de 60% des catalogues de la BnF étaient intégrés dans l'application. L'objectif fixé pour 2015 est d'atteindre 80 %. L'exhaustivité n'est pas visée à terme car seules les données de bonne qualité ont vocation à être exposées.

Ce projet entre dans la stratégie globale de la BnF sur la période 2014-2016 exposée dans son contrat de performance²⁰¹ qui consiste à mettre l'accent sur les applications de consultation plutôt que sur les applications de production des données. Celles-ci devraient devenir la priorité du prochain contrat de performance. On retrouve sans surprise les grands axes stratégiques adoptés dans le cadre de la transition bibliographique : priorité à l'exposition des données devant la refonte des modes de production qui viendra dans un second temps.

Techniquement, les pages de data.bnf.fr sont exposées en RDF/XML ou en JSON. Si l'on est curieux on peut cliquer sur un bouton au bas de chaque page afin de visualiser le résultat en RDF. Les pages sont également enrichies avec des données conformes à RDF/a et à schema.org.

Un dump de data.bnf.fr est disponible, ainsi que des dumps plus petits, prédécoupés par types de données, en fonction des usages supposés ou connus. Par exemple l'ensemble des auteurs-personnes, des auteurs-organisations, des sujets, des lieux... Par ailleurs un service d'interrogation des données en SPARQL a été ouvert en septembre 2014²⁰². Des exemples de requêtes sont proposées pour aider l'utilisateur à se familiariser avec ce langage²⁰³.

Si l'ensemble des données bibliographiques de la BnF est placé sous la licence ouverte de l'État depuis 2014, les données RDF de data.bnf.fr le sont depuis l'ouverture du service en 2011.

Les URI pérennes utilisées par la BnF sont formées à partir d'identifiants ARK²⁰⁴ (*Archival Resource Key*).

La BnF a recours autant que possible à des vocabulaires existants mais contrairement à l'ABES, face à certains besoins spécifiques elle a également créé une ontologie propre, bnf-onto²⁰⁵.

Les données de data.bnf.fr sont liées avec d'autres jeux de données : Geonames pour les lieux, VIAF ou Dbpedia pour les personnes. Les alignements concernant les sujets sont encore plus riches à partir de Rameau : Agrovoc (le vocabulaire contrôlé de la FAO, l'Organisation des Nations Unies pour l'alimentation et l'agriculture), le Thésaurus W des Archives nationales de France ou encore les référentiels de la *Deutsche Nationalbibliothek* et de la Bibliothèque du Congrès.

²⁰⁰ Sur la spécificité des interfaces en relation avec le web sémantique, voir le mémoire de Suzanne Goupil, *Ergonomie des interfaces du web sémantique : une nouvelle donne ? Les exemples de data.bnf.fr et du Centre Pompidou virtuel*, mémoire de Master Design d'Interface Multimédia et Internet sous la direction de Gildas Illien, 2014

²⁰¹ BnF, *Contrat de performance 2014-2016*, op.cit.

²⁰² <http://data.bnf.fr/sparql>

²⁰³ http://data.bnf.fr/docs/doc_requetes_data.pdf

²⁰⁴ ARK est un système d'identifiants pérennes administré par la California Digital Library. <https://wiki.ucop.edu/display/Curation/ARK>

²⁰⁵ <http://data.bnf.fr/ontology>

Quelles réutilisations ?

Nous l'avons déjà évoqué, les moteurs de recherche sont d'une certaine façon les premiers utilisateurs des données exposées. Les pages de data.bnf.fr ressortent en bonne position dans les résultats lorsqu'elles concernent des ressources rares : « Lorsqu'il y a peu de concurrence (par exemple en histoire des sciences ou des corpus juridiques, généralement peu visibles), le référencement est dopé. Pour les corpus déjà fortement disponibles et visibles ailleurs (ex : littérature), l'impact est positif, mais moins manifeste. Globalement, la sémantisation des pages favorise largement la longue traîne des œuvres peu disponibles et peu consultées²⁰⁶. »

Tout comme l'ABES, la BnF est confrontée à la difficulté de connaître les réutilisateurs de ses données. Les seules informations disponibles sont les statistiques de téléchargement des dumps, sans que cela en apprenne davantage sur les usages. La BnF communique notamment par le biais de la page dédiée data.bnf.fr²⁰⁷.

Citons quelques exemples de réutilisations. IfVerso²⁰⁸, la plate-forme du livre traduit de l'Institut français, utilise les données de data.bnf.fr pour regrouper des traductions autour d'une même œuvre. Isidore, portail de recherche en sciences humaines et sociales, s'appuie sur les données de Rameau exposées en RDF.

La BnF utilise elle aussi cet outil pour des besoins internes. Data.bnf.fr est en effet un outil de retraitement des données et d'amélioration du catalogue, celui-ci étant constitué de différentes strates comportant des niveaux de qualité divers. Nous reviendrons un peu plus loin sur l'utilisation de data.bnf.fr comme outil de FRBRisation du catalogue.

Enfin soulignons que data.bnf.fr est aussi un outil de marketing, une vitrine servant à attirer les visiteurs. La BnF communique beaucoup autour de ce projet, ce qui lui permet aussi de valoriser l'investissement public représenté par le catalogue en montrant son utilité au-delà de sa fonction initiale.

Bilan

Les deux agences exposent toutes deux leurs données bibliographiques sur le web de données mais selon des approches un peu différentes. Alors que data.bnf.fr est un nouveau produit, voire une marque, qui agrège les applications existantes, avec une interface spécifique, l'ABES fait le choix de mettre en avant ses applications existantes (IdRef, Sudoc...). Data.bnf.fr est de fait plus orienté vers le grand public tandis que pour l'instant l'ABES s'adresse davantage à des développeurs, notamment en leur proposant des services autour des données. Si l'on compare avec les bibliothèques nationales à l'étranger, la plupart axent leur stratégie sur la diffusion de données brutes comme le fait l'ABES. Par exemple la *British Library* propose avec la *Linked Open British National Bibliography* une diffusion gratuite de ses données en licence *Creative Commons CC0* (donc sans imposer de mention de source), accessibles par dump ou par SPARQL²⁰⁹. La Bibliothèque nationale d'Espagne propose quant à elle une application proche de celle de

²⁰⁶ WENZ Romain, « Data.bnf.fr : la fabrique », in *Synthèse de la journée d'étude du GFII* : « Données culturelles et Linked Open Data : valoriser le patrimoine public dans le web des données », 26 mars 2013, Maison de l'Europe. Disponible sur <http://www.gfii.fr/fr/document/journee-d-etude-du-gfii-donnees-culturelles-et-linked-open-data-valoriser-le-patrimoine-public-dans-le-web-de-donnees>

²⁰⁷ <http://data.bnf.fr/about>

²⁰⁸ <http://www.ifverso.com/>

²⁰⁹ <http://bnb.data.bl.uk/>

la BnF, datos.bne.es²¹⁰ rassemblant des ressources autour de pages œuvres, auteurs et thèmes.

Interrogation croisée des données

Dans le cadre de leurs réflexions sur la mutualisation de leurs données, les deux agences envisagent la possibilité d'offrir des services permettant d'effectuer des requêtes croisées sur leurs entrepôts de données respectifs. Cela n'est techniquement pas encore possible ni pertinent, car le SPARQL de l'ABES n'est pas prêt tandis que celui de la BnF ne couvre pas l'exhaustivité de ses données. Mais pouvoir interroger toutes les données des agences simultanément est sans aucun doute un objectif intéressant, car si elles contiennent nécessairement des redondances elles sont aussi complémentaires, compte tenu de leurs champs d'action respectifs.

Des cas d'utilisation réels issus de demandes de chercheurs permettent d'illustrer en quoi ce type de service serait utile. Ces cas sont exposés dans un article de Gildas Illien pour lequel les catalogues ont vocation à devenir des sources de *data mining* : « le catalogue de demain ne sera plus seulement un outil de recherche et de localisation d'information ; il faut aussi l'imaginer comme une vaste base de données à part entière que les chercheurs auront envie de faire parler²¹¹. » Le premier cas est celui d'un chercheur en sociologie s'intéressant aux journalistes écrivant des livres. Les autorités du catalogue général contiennent à cet égard une mine d'information mais qui sont difficiles à exploiter compte tenu de la rigidité des formats MARC. Il en est de même pour le second cas, une doctorante souhaitant obtenir la liste classée par nationalité des auteurs nés en Afrique et publiés en France.

Dans les deux cas l'interrogation croisée des données de l'ABES et de la BnF augmenterait la pertinence des réponses, en fournissant en une seule fois une information sur l'état des données bibliographiques possédées en France selon tel ou tel critère. Si l'on ne peut éviter de penser que la mise en place d'un dispositif d'interrogation croisée pourrait être évitée s'il existait un catalogue national, on ne peut que souligner l'intérêt d'une telle possibilité, si elle était offerte au public accompagnée d'une documentation d'aide à la formulation de requêtes, à la manière de ce que propose déjà data.bnf.fr. On peut en effet penser qu'un tel service serait d'une grande utilité au moins pour des chercheurs souhaitant explorer les possibilités du *data mining*.

MUTUALISER LES AUTORITÉS

L'idée évoquée précédemment selon laquelle le CBU continue à être un idéal d'actualité repose sur la conviction que deux éléments du web sémantique sont tout particulièrement essentiels pour atteindre cet objectif : « la maintenance d'identifiants uniques, normalisés, pérennes et actionnables » qui sont « des passerelles sans lesquelles on ne peut rien localiser, rien lier, rien construire²¹² » et les référentiels d'autorité, véritable « trésor de guerre » des bibliothèques. Le W3C

²¹⁰ <http://datos.bne.es/>

²¹¹ ILLIEN Gildas, « Metadata mining : fouiller les données des catalogues », in *Arabesques*, n°76, octobre 2014.

Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

²¹² BOURDON Françoise, ILLIEN Gildas, *op.cit.*

va dans le même sens : « Les données d'autorité de bibliothèque qui se rapportent aux noms et aux sujets contribueront à réduire la redondance existant sur le web au sein des descriptions bibliographiques grâce à l'identification claire d'entités clés qui seront partagées sur le web de données.²¹³»

Les bibliothèques manipulent trois grands types de référentiels d'autorités qui comportent chacun des enjeux propres : les personnes et collectivités, les sujets et enfin les titres et œuvres. Voyons désormais comment l'ABES et la BnF collaborent autour de ces référentiels.

Les personnes et collectivités

L'identification des personnes et collectivités, en particulier des auteurs, est fondamentale à la fois pour des questions liées à la gestion des droits, comme le montre le rapport Lescure²¹⁴, mais aussi à la bibliométrie et l'évaluation de la recherche, enjeu de plus en plus important pour les chercheurs et leurs institutions de rattachement. Cette identification est rendue possible par un écosystème de référentiels qui ont de plus en plus vocation à être interopérables les uns avec les autres.

Contexte : des référentiels internationaux

Revenons un instant au contrôle bibliographique universel. La gestion des autorités dans le CBU consistait avant tout à rechercher une uniformisation des noms d'auteur, de manière à accréditer une forme acceptée au niveau international pour choisir, par exemple, entre « DH Lawrence », « David Herbert Lawrence » ou encore « DH (David Herbert) Lawrence ». Gordon Dunsire et Mirna Willer²¹⁵ expliquent que cet objectif se modifie à partir de 2001, lorsqu'une nouvelle approche voit le jour à l'IFLA avec la publication des *Guidelines for authority records and references*²¹⁶, dans lesquelles on voit le terme « *authorized* » supplanter le terme « *uniform* ». Le critère important devient le fait qu'une agence bibliographique nationale valide la forme d'un nom d'auteur (forme qui fait alors autorité). Si les agences valident des formes différentes, cela n'a pas d'importance puisque l'on peut désormais les lier entre elles, c'est-à-dire les déclarer équivalentes, les « aligner ». Cette nouvelle conception du contrôle d'autorité est illustrée par la création du référentiel VIAF (*Virtual International Authority File*) en 2003.

VIAF

Les membres fondateurs de VIAF sont OCLC, la *Deutsche Nationalbibliothek* et la Bibliothèque du Congrès, rejoints par la BnF en 2007. Après une phase d'expérimentation, VIAF est lancé en 2011 en tant que produit d'OCLC. Il compte aujourd'hui 43 membres, principalement des bibliothèques nationales et des catalogues collectifs. IdRef, la base des autorités du Sudoc, fait partie de VIAF depuis 2012. La gouvernance repose sur un conseil de membres (*VIAF Council*) qui délibère et peut formuler des demandes d'évolution à OCLC.

Le principe du fonctionnement de VIAF repose sur l'agrégation des données de chaque partenaire : aucune donnée n'est produite. Les différentes notices sont

²¹³ GROUPE D'INCUBATION W3C, *op.cit.*

²¹⁴ LESCURE Pierre, *op.cit.*

²¹⁵ WILLER Mirna, DUNSIRE Gordon, *Bibliographic information organization in the semantic web*, Oxford : Chandos Publishing, 2013.

²¹⁶ IFLA Working Group on GARE Revision, *Guidelines for Authority Records and References*, München, K.G. Saur, 2001. Disponible sur <http://www.ifla.org/files/assets/hq/publications/series/23.pdf>

rapprochées pour créer ce que l'on appelle une grappe (ou *cluster*), à partir d'algorithmes visant à déterminer s'il s'agit bien de la même autorité. Ces algorithmes reposent sur différents critères qui sont pondérés. Ainsi les chiffres et les identifiants sont plus importants que les simples chaînes de caractères. En cas de doute sur le rapprochement des données, VIAF crée une grappe supplémentaire. Il s'agit donc en confrontant les données de mettre en évidence les convergences et divergences constatées et de laisser l'utilisateur en tirer des conclusions. On peut donc dire que VIAF reste neutre vis-à-vis des règles de catalogage des différents contributeurs²¹⁷.

La qualité et la valeur ajoutée de VIAF reposent par conséquent sur la qualité des données des partenaires. « Si les données entrantes sont suffisantes et de grande qualité, le résultat de l'appariement est fiable. Si les données fournies sont de mauvaise qualité, pauvres ou indifférenciées, le travail de rapprochement est compromis et le risque est élevé de créer de nouvelles grappes en doublon pour la même entité²¹⁸ ».

Un identifiant pérenne est attribué au niveau de chaque grappe VIAF générée. Les algorithmes étant rejoués continuellement, les grappes sont dans les faits mises à jour au fur et à mesure des imports de données envoyées par les partenaires.

VIAF est accessible gratuitement sur le web²¹⁹. L'interface de consultation permet de visualiser pour une grappe donnée chacun des partenaires contributeurs, matérialisé par son drapeau national (BnF) ou son logo (Sudoc). Les données brutes de VIAF sont disponibles et réutilisables dans divers formats (MARC, XML, RDF...) et sont placées sous la licence ODC-BY.

VIAF intègre peu à peu de nouveaux types de partenaires, tels que Getty Institute et Wikipedia. Le fichier évolue aussi afin de prendre en compte d'autres types de données d'autorité, telles que les collectivités mais aussi les œuvres ou les noms géographiques.

ISNI

L'ISNI (*International Standard Name Identifier*) repose quant à lui sur une norme ISO (27729) publiée en 2012. Au moment de l'élaboration de l'ISNI, un autre identifiant était à l'étude dans la communauté des bibliothèques, l'ISADN (*International Standard Authority Data Number*) mais le groupe de travail de l'IFLA dédié à cette question²²⁰ a recommandé en 2008 de ne pas créer de nouvel identifiant et d'utiliser l'ISNI.

Celui-ci est un identifiant multi-domaines « s'appliquant aux personnes et aux collectivités impliquées dans les contenus créatifs²²¹ ». Il peut s'agir par exemple des auteurs, éditeurs, traducteurs, illustrateurs mais aussi de sociétés de gestion de droits ou d'organismes de recherche.

La gouvernance de l'ISNI repose sur le dispositif suivant : l'agence internationale ISNI (ISNI-IA) est l'autorité d'enregistrement ISO de l'ISNI. Parmi ses membres se trouvent la CENL (*Conference of European National Librarians*) représentée par la *British Library* et la BnF qui sont en charge du contrôle qualité

²¹⁷ ANGJELI Anila, MAC EWAN Andrew, BOULET Vincent, « ISNI et VIAF transforment le paysage : pour des identités fiables et solides », in 80^e Congrès de l'IFLA, Session 86 - *Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?*, Lyon, 18 août 2014. Disponible sur <http://library.ifla.org/985/>

²¹⁸ *ibidem*

²¹⁹ <http://viaf.org/>

²²⁰ FRANAR (*Functional Requirements and Numbering of Authority Records*)

²²¹ ANGJELI Anila, MAC EWAN Andrew, BOULET Vincent, *op.cit.*

de la base. Le rôle d'agence d'attribution de l'ISNI (ISNI-AA) est quant à lui tenu par OCLC, qui a la charge de la gestion de la base de données. Puis le réseau est composé d'agences d'enregistrements et de simples membres.

Il faut noter que l'ISNI est un identifiant pour les identités publiques, ce qui signifie qu'un auteur dont l'identité est connue mais qui publie aussi sous un pseudonyme est susceptible de se voir attribuer plusieurs ISNI. Ainsi un ISNI a été attribué à John le Carré²²² et un autre à David Cornwell²²³.

ISNI s'appuie sur des données d'environ 70 contributeurs, dont ceux de VIAF. En 2011 au moment de sa constitution, les données de VIAF ont été utilisées comme fichier de départ. Mais contrairement à VIAF des interventions sur les données peuvent être réalisées via le contrôle qualité ou suite à des demandes du public. La base est accessible gratuitement sur le web²²⁴ et les données exposées en RDF. Un identifiant ISNI se matérialise sous la forme d'une URI pérenne.

Bien que VIAF contribue à l'ISNI, l'objectif et la méthode de l'ISNI diffèrent profondément de ceux de VIAF. « Il a une politique stricte qui découle de son objectif premier : consolider chaque identité publique et attribuer à chacune un identifiant ISNI unique et fiable qui puisse dès lors être diffusé largement et cité en toute confiance²²⁵. »

L'ISNI n'a pas vocation à remplacer les autres identifiants, ceux-ci ayant des fonctions propres dans les systèmes d'information locaux. Mais il a pour but de faire passerelle avec ceux-ci, en ajoutant une couche d'interopérabilité entre tous les identifiants, ce qui lui confère un fort potentiel dans l'environnement du web sémantique pour lier les données des différentes communautés.

La BnF est devenue en 2014 l'une des agences d'enregistrement ISNI. Sa politique consiste pour le moment à attribuer des identifiants ISNI aux contributeurs des documents présents dans ses collections. Mais une perspective d'évolution se dessine à la fin de l'année 2014. Elle consisterait à ce que d'autres acteurs participent avec la BnF à une fédération d'agences ISNI, chacune ayant la responsabilité de l'attribution d'identifiants pour son périmètre, tandis que la BnF continuerait à être l'interlocuteur de l'agence internationale. L'un de ces acteurs est l'ABES, qui prendrait ainsi en charge l'identification des doctorants, des enseignants chercheurs, des laboratoires et diverses structures de l'enseignement supérieur et de la recherche. Les Archives nationales souhaitent également participer à un tel système, elles auraient alors en charge en particulier l'identification des collectivités. D'autres pistes d'élargissement en direction de l'interprofession du livre sont à l'étude. A compter du 1^{er} janvier 2015, l'ABES devient membre de l'ISNI pour une durée d'un an, en attendant que la fédération se construise sur la plan national²²⁶.

ORCID

ORCID (*Open Research and Contributor Identifier*) est aussi une initiative internationale lancée en 2012 afin d'attribuer un identifiant unique aux auteurs et contributeurs de l'enseignement supérieur et de la recherche. Cet identifiant a rapidement rencontré un certain succès auprès des chercheurs. Mais contrairement à l'ISNI il est relativement peu contrôlé, les chercheurs pouvant créer ou modifier eux-mêmes les données les concernant. D'autre part il ne gère pas les collectivités. Un accord

²²² <http://isni.org/isni/000000012284048X>

²²³ <http://isni.org/isni/0000000037129212>

²²⁴ <http://www.isni.org/>

²²⁵ ANGJELI Anila, MAC EWAN Andrew, BOULET Vincent, *op.cit.*

²²⁶ <http://fil.abes.fr/2014/12/18/labes-bientot-membre-de-lisni/>

d'interopérabilité a été conclu avec l'ISNI. ORCID utilise d'ailleurs l'ISNI pour identifier les institutions auxquelles appartiennent les chercheurs.

Alignement des identifiants ABES et BnF

Le web de données doit permettre à différents référentiels de se lier entre eux. Dans le cadre de leur coopération, l'alignement des identifiants personnes et collectivités de l'ABES (IdRef) et de la BnF (ISNI) est un chantier fondamental. Il s'agit d'ailleurs de celui qui est actuellement le plus avancé, différents processus étant déjà à l'œuvre pour permettre de réaliser ces liens.

Tout d'abord, certains alignements existent déjà de fait, dans la mesure où le Sudoc est en partie alimenté par les données d'autorité de la BnF. Les notices IdRef conservent ainsi l'identifiant d'origine de la BnF (numéro FRBNF), ce qui permet de les déclarer équivalentes. Ce cas permet de souligner à quel point il est important de disposer de données de bonne qualité dans l'environnement du web de données. L'ABES a ainsi dû ouvrir un chantier de correction de données lié à des doublons de numéros FRBNF, que les catalogueurs oublient parfois de supprimer lorsqu'une notice est utilisée comme base pour en créer une nouvelle. On voit que la qualité du catalogage doit continuer à être valorisée y compris dans le cadre de processus de récupération de notices extérieures. L'ABES s'emploie à faire prendre conscience de ce fait : « La question des identifiants [...] est de plus en plus présente dans les priorités de l'ABES car ils interviennent dans les alignements de « nos » données avec celles d'autres catalogues, à commencer par nos applications STAR, STEP ou Calames et plus loin BnF, VIAF, ISNI. En effet, les identifiants contenus dans les notices sont bien souvent l'unique clé permettant à une application extérieure de venir retrouver les informations des autorités du Sudoc²²⁷ ».

D'autre part, des alignements sont réalisés par l'intermédiaire de VIAF, lorsque les identifiants BnF et ABES se retrouvent agrégés dans une même grappe. Ainsi, lorsque la BnF charge à son tour les données de VIAF, elle récupère également les identifiants IdRef, qu'elle expose dans data.bnf.fr. C'est donc VIAF, un produit OCLC, qui a fourni aux deux référentiels d'autorités des agences bibliographiques françaises l'opportunité de se rencontrer sur le web de données, un détour qui peut paraître pour le moins étonnant.

Enfin les données des deux agences se rencontrent également par le biais de l'ISNI. Tout d'abord parce que la BnF diffuse désormais les ISNI dans ses produits bibliographiques. Ceux-ci sont donc mis à la disposition de l'ABES qui va les intégrer dans le Sudoc à partir de 2015. Elle récupérera les ISNI correspondant à des notices de son catalogue et ne figurant pas dans les produits de la BnF par l'intermédiaire de VIAF. Enfin, l'attribution d'ISNI directement par l'ABES dès 2015 devrait contribuer à cette généralisation d'un identifiant unique et interopérable.

On va donc de plus en plus disposer dans les deux catalogues d'un même identifiant pour les mêmes contributeurs. Soulignons à nouveau que cela ne signifie pas la fin des identifiants locaux. Ainsi l'ABES va continuer à utiliser IdRef, qui doit également pouvoir communiquer avec les identifiants du monde de la recherche. ORCID a été mentionné, mais IdHAL, l'identifiant de la plate-forme d'archives ouvertes HAL²²⁸ avec lequel IdRef est également lié, en est un autre

²²⁷ <http://pункtokomo.abes.fr/2014/10/28/un-chantier-qualite-sur-les-n-frbnf-multiples/>

²²⁸ <https://hal.archives-ouvertes.fr/>

exemple. En effet « tandis qu'un identifiant unique au sein d'un référentiel unique est voué à l'incomplétude, à l'isolement et/ou à l'obsolescence, l'interopérabilité et l'alignement des différents référentiels ouvrent des perspectives nouvelles pour la qualité des catalogues et pour l'offre de services aux usagers²²⁹. »

La gestion des autorités personnes comporte une spécificité qui doit être mentionnée, celle de la question des données personnelles qui sont manipulées et que l'exposition sur le web de données rend plus visibles. On touche ici également à des questions d'« e-réputation ». Les personnes constatant des erreurs sur les informations les concernant, ou ayant un repentir lié à une œuvre ou encore ne souhaitant pas voir divulguer leur date de naissance peuvent faire une demande de correction à la BnF. L'ABES en reçoit elle aussi de plus en plus dans le cadre du signalement des thèses. La BnF refuse de supprimer ces informations, afin de ne pas déroger à sa mission de dépôt légal. Elle accepte en revanche de masquer certaines données à l'affichage. Les deux agences ont décidé de saisir conjointement la CNIL (Commission nationale de l'informatique et des libertés) afin de vérifier que cette exposition croissante de données personnelles à l'extérieur des catalogues ne doit pas conduire à une évolution des procédures actuellement en vigueur.

Les sujets

La question des autorités sujets n'est pas tant celle de la mutualisation, puisque comme on l'a vu il existe déjà un fichier national d'autorités sujets sous la forme de Rameau, mais plutôt celle de la valorisation de ce référentiel sur le web sémantique.

En 2008 Rameau a été traduit en SKOS, formalisme permettant de représenter les thésaurus ou autres langages contrôlés en RDF, par un chercheur de l'Université libre d'Amsterdam. Ce prototype a été repris par la BnF en 2011 et versé dans data.bnf.fr. Il est depuis récupérable sous forme de dump et est utilisé pour construire les pages thématiques de data.bnf.fr et permet lors de la navigation de rebondir vers une autre notice Rameau ou un autre type de notice d'autorité. Il est par ailleurs aligné avec d'autres référentiels sujets comme cela a déjà été évoqué. L'ABES le valorise également dans le cadre de l'exposition d'IdRef, qui rappelons-le contient tout type d'autorité.

Rameau comporte de nombreux atouts pour le web de données. Outre les qualités propres aux référentiels d'autorités en général (données structurées, contrôlées, liées) Rameau assure « une couverture encyclopédique des domaines du savoir en phase avec la réalité de la production éditoriale et des besoins documentaires » et s'appuie sur « un réseau de contributeurs organisé et actif [...] qui repose sur des professionnels qualifiés et dispose d'instances de coordination et de validation garants de la qualité et de la cohérence des données diffusées ; une communauté d'utilisateurs qui, par sa diversité tant institutionnelle que géographique, constitue un gage d'interopérabilité entre des pratiques et des systèmes différents²³⁰. »

Rameau comporte un fort potentiel d'utilisations. Un exemple de cas d'application est celui du multilinguisme. Le projet MACS (*Multilingual access to subject*) est un programme européen visant à créer des liens entre Rameau, le référentiel des autorités en allemand de la DNB, SWD (*Schlagwortnormdatei*) et la LCSH en anglais. L'objectif est de permettre à un utilisateur d'interroger dans sa propre langue les catalogues des bibliothèques en langue étrangère. Ce programme a abouti à la réalisation d'un prototype

²²⁹ MISTRAL François, « Articulation ISNI-IdRef : un enjeu pour l'identification pérenne », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

²³⁰ Comité d'orientation Rameau, *op.cit.*

maintenu par TEL (*The European Library*). Mais il s'appuie sur une technologie aujourd'hui obsolète, des traitements manuels coûteux, et la décision a été prise en août 2014 de mettre fin au projet. Les acquis de ce projet seront vraisemblablement repris afin de les valoriser sur le web de données, selon des modalités qui restent à déterminer.

De leur côté l'ABES et la BnF ont pris acte, lors de leur réunion opérationnelle de février 2014, du fait qu'il serait opportun d'avancer ensemble sur la modernisation de Rameau et son adaptation à l'environnement du web sémantique. Néanmoins pour le moment aucun projet concret ne se dessine ; il faut probablement souligner qu'un référentiel sujets, par rapport à un référentiel de personnes par exemple, reste très complexe à manipuler. Ce chantier n'est donc pas prioritaire dans l'immédiat, ni dans la collaboration entre les deux agences, ni même chez un acteur tel qu'OCLC.

Les titres et œuvres

Le modèle FRBR est structuré autour de la notion d'œuvre, mais il n'existe pas encore à proprement parler de fichier d'autorités « œuvres ». Il existe néanmoins des fichiers d'autorités « titres » qui peuvent aider à leur construction.

La FRBRisation des catalogue consiste en partie à créer des notices d'œuvres lorsque cela est possible. Il n'existe pas encore de système produisant des données selon le modèle FRBR, celle-ci est donc réalisée a posteriori sur les données des catalogues. Le résultat est donc relativement imparfait du point de vue du modèle FRBR. La BnF parle de « semi-FRBRisation ». Il s'agit d'exploiter les données du catalogue afin d'une part de les exposer selon le modèle FRBR comme cela est fait dans data.bnf.fr, d'autre part d'enrichir les liens du catalogue afin de préparer les données à une vraie FRBRisation.

Une autre manière de s'y préparer consiste à mettre l'accent sur le respect des pratiques de catalogage. Comme l'explique Philippe Le Pape, « l'une des difficultés à surmonter au cours du processus du passage aux FRBR — et non la moindre — sera la mise à niveau des données rétrospectives, qu'il faudra «FRBRiser» à coup de procédures informatiques reposant sur des algorithmes de conversion²³¹.» Ces algorithmes chercheront à identifier les œuvres, les manifestations, les expressions, sur la base des données figurant dans le catalogue. D'où l'importance de la qualité de ces données, sur laquelle il est possible d'être vigilant dès aujourd'hui au moment de leur production.

Ces deux approches peuvent se combiner. Comparons brièvement les méthodes respectives de la BnF et de l'ABES.

Les projets de la BnF l'incitent tout particulièrement à manipuler la notion d'œuvre. C'est le cas avec le dépôt légal des *e-books* qui va avoir pour conséquence la gestion d'un nombre croissant de manifestations d'une même œuvre (version papier et versions numériques dans divers formats).

L'outil principal de FRBRisation du catalogue est l'application data.bnf.fr, via la mise au point de différents types d'algorithmes : algorithmes d'alignement visant à créer des liens entre des notices bibliographiques et des notices de titres

²³¹ <http://rda.abes.fr/2012/04/20/preparer-la-frbrisation-des-donnees/>

existantes, mais aussi un algorithme de regroupement, qui agrège des manifestations en l'absence de notice de titre afin de générer une œuvre. Comme cela a été évoqué, ces traitements sont ensuite contrôlés par des catalogueurs qui aident à déterminer le seuil de confiance de ces algorithmes.

Une autre piste de FRBRisation des données à la BnF se présente dans le cadre de la création du registre ReLire. Pour ce projet, ARROW (*Accessible Registries of Rights Information and Orphan Works towards Europeana*) et TEL procèdent au regroupement des différentes manifestations autour des œuvres. Compte tenu de l'objet du registre, ce travail s'effectue sur un corpus d'œuvres du XXe siècle. Les équipes de ReLire et de data.bnf.fr vont collaborer pour comparer les résultats de leurs traitements respectifs sur ce même corpus, ce qui permettra d'évaluer en retour les algorithmes utilisés de part et d'autre. La BnF est donc dans une démarche expérimentale, les résultats de ces traitements seront d'ailleurs testés dans le cadre de la refonte de l'interface du catalogue général, celle-ci visant à proposer une expérience de navigation FRBRisée (regroupement des manifestations autour des œuvres).

La BnF a fait le choix pour l'instant de ne pas faire évoluer ses consignes de catalogage, conformément à sa stratégie de traiter l'exposition des données avant leur production.

L'ABES a adopté une démarche différente. Elle avait comme objectif de préparer la FRBRisation en insistant sur le respect des consignes de catalogage, diffusant ce qui a été appelé la « première vague » de consignes pour le Sudoc. Il s'agissait par exemple de transcrire systématiquement le titre original de l'œuvre, afin de pouvoir au moment de la FRBRisation rapprocher les différentes manifestations autour de l'œuvre ; identifier la langue originale et la langue de publication, afin de pouvoir discerner différentes expressions ; transcrire systématiquement les identifiants normalisés qui peuvent aider à la création de liens (ISBN, ISSN...) ; conserver les identifiants de notices dérivées, etc²³².

La perspective du changement d'outil avec le SGBM incitait par ailleurs à une certaine prudence. Parallèlement, des tests sur des algorithmes ont été menés afin de modifier les données a posteriori. Mais pour des raisons diverses ces deux pistes ont été mises en suspens et il n'y a pas eu de deuxième vague de consignes à destination du réseau.

Une autre voie est aujourd'hui étudiée par l'ABES avec l'aide de son prestataire OCLC. Il s'agit d'utiliser un outil qui compare les notices et effectue des regroupements afin de générer une notice d'œuvre. Cet outil a été utilisé dans le réseau des bibliothèques des Pays-Bas qui est doté d'un environnement CBS similaire à celui de l'ABES. Ces algorithmes sont assez proches de ceux qui sont utilisés par ailleurs par OCLC pour générer des notices d'œuvres dans VIAF.

Les deux agences expérimentent donc des processus de FRBRisation chacune de leur côté, tout en s'informant mutuellement de leurs avancées respectives, voire envisagent de partager leurs algorithmes pour les tester sur leurs propres données. Il s'agit donc d'expérimenter et de comparer les méthodes, avant de penser à la mutualisation. Mais celle-ci est jugée nécessaire car les données d'autorité œuvres sont fondamentales et difficiles à produire sur des volumes importants. L'objectif à terme est bien la constitution d'une base d'œuvres au niveau national. Les deux agences

²³² <http://rda.abes.fr/2012/04/20/preparer-la-frbrisation-des-donnees/>

envisagent déjà, quand elles seront prêtes, une répartition pour la création d'œuvres selon leurs domaines de spécialité respectifs (domaine français et documents spécialisés pour la BnF, domaine étranger et thèses pour l'ABES.) L'objectif reste de FRBRiser les données existantes avant la prochaine étape, la production de données conformes au modèle FRBR.

Une question reste en suspens concernant les identifiants à assigner aux œuvres. À l'heure actuelle on peut mentionner l'existence de l'ISTC²³³, (*International Standard Text Code*), un identifiant normalisé mais qui s'applique uniquement aux œuvres textuelles, au niveau de l'expression et non de l'œuvre, et qui du point de vue des agences bibliographiques comporte l'inconvénient d'être géré en France exclusivement par Électre. Une autre piste se dessine sous la forme d'un identifiant propre à OCLC, attribué aux œuvres dans VIAF. La question de l'interaction des œuvres VIAF avec la production d'œuvres ABES et BnF va également devoir être posée, peut-être sur un modèle similaire aux autorités personnes.

Vers un fichier national d'autorités ?

Les équipes de l'ABES et de l'IBN s'emploient à imaginer de manière prospective des scénarios de mutualisation de tous les types d'autorités qui viennent d'être évoqués. L'idée qui émerge est celle d'une base d'autorités communes, qui servirait de pivot aux bases de production respectives des deux agences, voire à celles d'autres contributeurs. Ce fonctionnement est celui d'IdRef qui est la base pivot pour Calames, le Sudoc et les thèses. Le fait d'avoir des autorités gérées en dehors des systèmes de production serait tout à fait conforme à la tendance des autorités à devenir de plus en plus autonomes et permettrait de faciliter les alignements avec des référentiels d'autres communautés.

Cette hypothèse s'inscrit dans un contexte particulièrement favorable compte tenu des perspectives de refonte des systèmes des deux agences. Le système de production de la BnF dans le cadre du prochain contrat de performance et le projet Sudoc II dans un délai probablement comparable. Restent à déterminer les modalités de mise à jour de cette base, comment elle se synchronisera avec les environnements de production de chacun, et surtout comment s'organisera sa gouvernance.

PRÉPARER LA TRANSITION BIBLIOGRAPHIQUE

Un nouveau cadre pour adopter RDA

Jusqu'à la fin de l'année 2014, les questions liées à l'adoption de RDA étaient traitées au sein d'un groupe nommé « RDA en France », chargé d'examiner l'opportunité et les modalités d'adoption de RDA, dans le cadre de la commission CG46 Documentation de l'AFNOR. Ce groupe se composait d'un groupe technique traitant des questions de normalisation, d'un groupe formation et d'un groupe SID/SIGB dédié aux systèmes d'information. Ce dispositif était piloté par le

²³³ « Défini par la norme ISO 21047:2009, l'ISTC est un identifiant unique, permanent et de valeur internationale pour les œuvres textuelles, c'est-à-dire les contenus constitués en majeure partie d'une combinaison de mots, indépendamment de leurs manifestations et des formes matérielles de leur publication. » http://www.bnf.fr/fr/professionnels/glossaire_numeros.html

groupe stratégique RDA en France composé de représentants de l'ABES et de la BnF mais aussi de l'AFNOR, d'associations professionnelles ou d'éditeurs. Ce cadre vient de changer, de deux façons différentes.

Tout d'abord, il a été décidé d'élargir officiellement le champ de travail de cette organisation. Il ne s'agit plus seulement de l'adoption de RDA mais de prendre en compte les autres aspects de la transition bibliographique, à savoir l'adoption progressive du modèle FRBR ou encore la diffusion des données bibliographiques sur le web, conformément à la stratégie adoptée en France. Il s'agit donc d'« inscrire la problématique de l'évolution des règles de catalogage dans un plus large mouvement d'évolution qui nécessite une démarche d'ensemble d'accompagnement au changement dans les établissements²³⁴. »

L'autre changement concerne la gouvernance de ce dispositif, qui est réduite à l'ABES et à la BnF. En effet le CSB (à savoir les deux établissements et leurs tutelles) prend la suite du comité stratégique RDA en France. L'organisation repose toujours sur trois groupes regroupant globalement les mêmes fonctions : formation, normalisation (ce groupe porte toujours le nom RDA en France) et systèmes et données. Les anciens partenaires du groupe stratégique siègent dans les trois groupes et continueront à être informés y compris dans d'autres instances telles que les journées d'associations professionnelles. Une instance assure la coordination technique des trois groupes, elle est composée des trois responsables de groupes et est nommée « Transition bibliographique », afin de rendre compte de l'élargissement de la problématique. Une instance de pilotage resserrée à la direction de l'ABES et de l'IBN rend compte au CSB. Ce nouveau dispositif disposera d'un site, transition-bibliographique.fr²³⁵, hébergé par l'ENSSIB comme l'était le site précédent RDA en France²³⁶.

Ainsi que nous l'avons déjà évoqué, RDA dans sa forme actuelle ne satisfait pas aux exigences françaises. Mais cela n'empêche pas qu'un important travail de normalisation continue. Il s'agit de réviser les normes de catalogage françaises en écrivant une norme dite de transition vers RDA, qui prenne en compte autant que possible les éléments de RDA. Cette norme est nommée *Transposition française de RDA* (RDA FR). Il ne sera pas nécessaire d'attendre que cette norme soit achevée pour en voir la publication, elle sera au contraire diffusée au fur et à mesure que des blocs cohérents seront achevés²³⁷. Ce sera par exemple le cas des éléments Forme du contenu, Type de médiation, Type de support qui intégreront notamment la zone 0 de l'ISBD²³⁸.

Outre la rédaction de la norme, un travail concret consistant à faire évoluer les pratiques de catalogage en émettant de nouvelles consignes est mené au fur et à mesure, dans le but de préparer la FRBRisation des catalogues. Ce travail est mené par le groupe technique en coordination avec le groupe formation et se traduit par l'alimentation du guide pratique du catalogueur de la BnF²³⁹ et du guide méthodologique Sudoc²⁴⁰.

La mise en pratique de ces nouvelles consignes est sans doute facilitée par l'existence d'un réseau de catalogage partagé tel que le Sudoc. Cela est moins évident pour les bibliothèques publiques, qui comme le souligne le rapport au CSB, « ne sont pas insérées dans un réseau national aussi formel et relèvent de collectivités locales très

²³⁴ ABES, BnF, *De RDA en France à la Transition bibliographique. Préconisations de l'ABES et de la BnF pour favoriser l'évolution des catalogues nationaux vers le web des données liées*, novembre 2014. Disponible sur http://www.bnf.fr/documents/1411_transition_bibliographique.pdf

²³⁵ <http://transition-bibliographique.fr/> Le site est en cours de construction au 17 décembre 2014

²³⁶ <http://rda-en-france.enssib.fr/>

²³⁷ <http://rda.abes.fr/2014/11/17/un-chapeau-en-forme-de-transition-bibliographique/>

²³⁸ La zone 0 (zone de la forme du contenu et du type de médiation) remplace l'indication générale du type de document. Pour plus d'informations voir <http://rda.abes.fr/2014/05/26/la-zone-zero/>

²³⁹ <http://guideducatalogueur.bnf.fr/>

²⁴⁰ <http://documentation.abes.fr/sudoc/>

diverses. Leurs problématiques se sont par ailleurs beaucoup éloignées des questions de catalogage, [...] ce qui rend plus difficile aujourd'hui l'adoption de bonnes pratiques sans compréhension des nouveaux enjeux²⁴¹». Cette difficulté à toucher les bibliothèques publiques rend d'autant plus importante la nécessité de convaincre leurs managers : le document présentant le nouveau dispositif met également en avant, outre la nécessité de former les catalogueurs, celle de sensibiliser les décideurs « aux enjeux stratégiques (maîtrise des données, organisation du travail, évolution des métiers)²⁴² ».

Le groupe formation a pour missions la sensibilisation et la formation aux enjeux des évolutions du catalogage : comprendre les limites des catalogues actuels, connaître le contexte normatif et les apports du modèle FRBR et du web sémantique. Il travaille en liaison avec les différents organismes de formation (Centres régionaux de formation aux carrières des bibliothèques, Mediadix...) et anime un réseau de formateurs. Voici quelques exemples d'intitulés de formations qui ont été dispensés en 2013-2014²⁴³ : *Évolution des catalogues et du catalogage ; Les catalogues au défi du Web. FRBR et RDA en France ; Mettre les catalogues de bibliothèques dans le Web de données.*

Le groupe systèmes et données a un rôle de veille sur les expériences de FRBRisation. Il assure aussi la communication auprès des administrateurs de SIGB des bibliothèques, notamment afin de leur donner des conseils en cas de réinformatisation, et auprès des éditeurs, pour les informer sur les évolutions du contexte normatif et les implications des nouvelles règles de catalogage.

Les travaux de ces groupes ont donc également une mission d'accompagnement du changement. Elle était déjà présente dans l'ancienne formule mais s'accroît encore avec l'élargissement officiel de la problématique. Attardons-nous un instant sur cet enjeu.

Conduire le changement

Nous proposons dans cette section quelques pistes pour imaginer ce que pourrait être un projet de conduite du changement²⁴⁴ à l'échelle nationale. L'ancien dispositif s'est jusqu'à présent concentré sur la communication autour de la normalisation en cours et sur la sensibilisation aux évolutions du catalogage, ce qui est déjà très important. Mais il serait intéressant d'appréhender le changement plus globalement, et d'envisager une coordination des actions à grande échelle, coordination qui pourrait être assurée en mode projet et qui pourrait être menée dans le cadre de ce nouveau dispositif.

Conduire le changement en mode projet permettrait une meilleure coordination et un meilleur pilotage des actions d'accompagnement menées. Cela comporterait en outre l'avantage, sur une durée limitée à définir, de mobiliser formellement des moyens, des ressources, certaines éventuellement dédiées au projet, plutôt que de confier ces tâches à des agents qui occupent par ailleurs d'autres fonctions et qui n'ont pas nécessairement la maîtrise du temps qu'ils

²⁴¹ ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique*, op.cit.

²⁴² *ibidem*

²⁴³ <http://rda-en-france.enssib.fr/Listedespr%C3%A9sentationsdonn%C3%A9es>

²⁴⁴ Nous utilisons ici le terme conduite du changement plutôt que celui d'accompagnement. Il s'agit d'une distinction subtile, mais certains théoriciens s'accordent sur le fait que l'accompagnement n'est qu'une partie de la conduite du changement, celle-ci incluant l'ensemble des actions à mener, et notamment une phase de diagnostic du changement.

peuvent y consacrer. La direction du dispositif a ici un rôle important de sensibilisation des tutelles à jouer afin d'obtenir ces moyens, ce qui peut s'avérer difficile car les sujets dont il est question sont assez techniques et il peut être complexe d'en expliquer les enjeux.

Caractériser le changement

Il est indispensable de mener préalablement à toute action un diagnostic du changement. Cette phase consiste à poser un certain nombre de questions, visant à déterminer l'ampleur et la nature du changement et à le caractériser. Par exemple, s'agit-il d'un changement technologique, organisationnel, culturel ? Nous l'avons vu, ces qualificatifs s'appliquent tous au contexte de la transition bibliographique. Donnons un nouvel exemple concernant le bouleversement culturel qu'implique l'exposition des données à la BnF. Dans le cadre du projet data.bnf.fr des notices dites de gestion sont désormais visibles. Elles sont intéressantes pour le web de données parce qu'elles comportent des liens mais en tant que telles, parce qu'elles ne correspondent pas aux canons traditionnels du catalogage, certains peuvent considérer qu'elles « salissent » le catalogue. Néanmoins cette fierté vis-à-vis des données est aussi un atout. En effet un diagnostic doit permettre de mettre en valeur les richesses à préserver lors d'un changement. En l'occurrence les compétences pointues des catalogueurs, leur système de valeurs, leur fierté liée à la richesse des données qu'ils produisent, sont de véritables atouts pour la réussite du web de données.

Le diagnostic doit permettre également de s'interroger sur la capacité à changer. Les agents des bibliothèques sont sur ce point relativement habitués au changement, étant données les évolutions technologiques permanentes auxquelles ils sont confrontés. Pour Émilie Liard (SCD Poitiers), formatrice à RDA dans le réseau Sudoc, la transition ne sera pas la plus difficile pour les catalogueurs, ceux-ci ayant l'habitude des normes. Elle insiste par contre sur le fait qu'il est important que les décideurs aussi soient familiarisés avec ces questions²⁴⁵.

Un autre point fondamental du diagnostic consiste en effet à déterminer précisément quelles sont les populations touchées et à quel degré. L'idéal est de pouvoir les représenter sur un schéma en forme de cible, afin de faire apparaître ces différents niveaux. Les critères à prendre en compte sont les différents métiers mais aussi les types de bibliothèques. Ainsi les catalogueurs de la BnF et ceux du Sudoc seraient vraisemblablement au centre de la cible, mais ceux des bibliothèques publiques également, pour des raisons différentes. Il faut en effet distinguer les catalogueurs qui vont continuer à cataloguer mais dont le métier va se transformer et ceux qui vont le faire de moins en moins et dont les attributions vont devoir évoluer vers des activités de médiation et d'accompagnement, comme le montre le rapport de l'IGB sur l'évolution des métiers²⁴⁶.

La question de la temporalité doit aussi être abordée : s'agit-il d'un changement à court ou long terme ? En l'occurrence, comme le souligne Françoise Leresche, il n'y aura « pas de grand soir », ces changements vont se faire de manière lente, progressive. Cette temporalité comporte ses avantages et ses inconvénients. Elle laisse le temps de s'adapter mais peut rendre les changements difficiles à comprendre. C'est le cas pour l'adoption de RDA : il en est question depuis un certain temps déjà, mais la concrétisation se fait attendre, ce qui peut susciter des incompréhensions ou un certain désintérêt. D'autre part les conséquences sont très différentes selon les situations locales,

²⁴⁵ LIARD Émilie, « RDA vu de l'intérieur », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 21 mai 2014. Disponible sur <http://canal2.u-strasbg.fr/video.asp?idvideo=12655>

²⁴⁶ AROT Dominique *et alli*, *op.cit.*

par exemple suite à la décision d'une bibliothèque publique de dériver ses notices de la BnF, ou lors de la mise en place d'un projet de *linked data* dans une bibliothèque spécialisée, ou par la modification des consignes de catalogage dans le Sudoc par vagues successives, ou encore lors d'un projet de réinformatisation avec un SIGB s'appuyant sur le modèle FRBR...

Cette phase sert aussi à déterminer les acteurs du projet. Ceux-ci doivent idéalement être issus de l'ensemble des publics touchés. Or jusqu'à présent, le groupe RDA en France dédié à la normalisation laissait percevoir un certain déséquilibre en faveur de la BnF, surreprésentée par rapport aux autres participants (ABES, SCD, bibliothèques spécialisées), celle-ci pourvoyant environ deux tiers des effectifs. Ce déséquilibre peut avoir tendance à renforcer l'idée que la normalisation est l'affaire de la BnF, alors qu'une plus grande représentation des autres acteurs serait également un levier pour le changement, ceux-ci pouvant constituer des relais dans les établissements. Mais cela nécessite de convaincre l'encadrement que participer à la normalisation comporte un intérêt pour l'établissement et que cela justifie le coût que cet investissement représente (en temps et frais de déplacement pour les établissements de province, les réunions ayant lieu à Paris).

Une fois le diagnostic du changement réalisé, il est possible d'identifier les leviers à actionner, et de cibler et dimensionner les ressources à affecter au projet. Dans le cadre de ce mémoire on peut d'ores et déjà supposer qu'il serait pertinent d'actionner au moins les trois leviers suivants : la communication, la formation et l'animation de réseaux de managers.

Informer, communiquer

Informer sur le contenu et les enjeux de la transition bibliographique revêt des difficultés. Il faut en effet souligner la complexité du sujet ainsi que l'imbrication de différents éléments qui peuvent en rendre la compréhension ardue. Pour illustrer ce point mentionnons que Philippe Le Pape avait intitulé son intervention lors des Journées ABES 2013, *Sudoc, un instant de perplexité*²⁴⁷, constatant celle du réseau face aux nouvelles consignes de catalogage, augmentée encore par d'autres sujets d'incertitude liés à l'évolution du Sudoc ou encore à l'adoption du code de catalogage RDA.

La profession de manière générale est mobilisée pour informer sur ces thématiques, que ce soit par les blogs de l'ABES mentionnés plus haut, les pages professionnelles de la BnF et au-delà via des blogs non institutionnels²⁴⁸ ou les réseaux sociaux. Il s'agit en outre d'un domaine en constante évolution qui s'écrit beaucoup au cours de journées d'études ou de conférences²⁴⁹.

Il y a donc une grande quantité de sources d'informations, mais chacune diffuse un message qui lui est propre. Il serait sans doute opportun de les coordonner. Le communiqué expliquant la stratégie française et présentant le nouveau dispositif répond en partie à ce besoin de clarification²⁵⁰. Mais comment celle-ci va être relayée ? Qui va effectivement lire ce document ?

Il y a peut-être une opportunité à saisir dans la refonte annoncée du site web. Le site RDA en France n'était pas jusqu'à présent un modèle de communication,

²⁴⁷ LE PAPE Philippe, « Sudoc, un instant de perplexité », [enregistrement vidéo], in Journées ABES 2013, Montpellier, 15 mai 2013. Disponible sur <http://canale2.u-strasbg.fr/video.asp?idvideo=11836>

²⁴⁸ Par exemple : BERMÈS Emmanuelle, Figoblog, [en ligne]. Disponible sur <http://www.figoblog.org/>

²⁴⁹ Voir par exemple la liste des manifestations constituant les sources de cette étude.

²⁵⁰ ABES, BnF, *De RDA en France à la Transition bibliographique, op.cit.*

étant plutôt destiné aux travaux des groupes. Mais le nouveau site sera peut-être plus ambitieux et pourrait permettre de donner une nouvelle visibilité aux enjeux de la transition bibliographique. Le regroupement des divers sujets sous un terme fédérateur est déjà une bonne base pour clarifier le contexte. Il serait sans doute utile que ce site se donne un objectif pédagogique, afin de devenir une source d'information de référence pour diffuser des messages cohérents et à jour, ce qui n'empêche pas de renvoyer sur d'autres sources d'informations, en les documentant, à la manière d'un portail. Il serait pertinent, outre le fond, de travailler particulièrement la forme et l'ergonomie du site, afin qu'il soit moderne et attractif, quitte à s'entourer ponctuellement de professionnels du design et de la communication écrite.

L'information ne doit pas se faire seulement à distance mais aussi en présentiel, selon des formules permettant aux agents de poser des questions et d'exprimer leurs craintes. Des réunions d'information sont proposées aux agents de la BnF, et on peut considérer que les journées ABES jouent aussi ce rôle. Cela est peut-être plus complexe à mettre en œuvre ou moins pertinent à l'échelle d'un établissement dans lequel peu de personnes sont touchées directement ou alors le seront sur le long terme. Dans ce cas, il faut souligner l'importance du rôle du management de proximité qui doit fournir en temps voulu un accompagnement adapté.

Former, gérer les compétences

La formation est l'un des leviers habituels de la conduite du changement. Elle peut s'appuyer avec profit sur une démarche de type GPEC (Gestion prévisionnelle des emplois et des compétences). La GPEC permet de définir les compétences à acquérir (compétences cibles) et de mettre en place des plans de formation pour passer des compétences initiales aux compétences cibles. Elle se décline en fonction de l'analyse effectuée lors du diagnostic qui prend en compte les différents métiers et niveaux d'implication de chacun.

Lors du dernier congrès de l'IFLA, les participants à l'atelier du groupe d'intérêt sur le web sémantique se sont interrogés sur le bon niveau de formation à dispenser en fonction de ces critères dans un projet de *linked data* en bibliothèque. Une formation technique poussée est nécessaire pour les personnels chargés du projet, mais une formation de base est également indispensable pour ceux qui peuvent intervenir ponctuellement. « Il s'agit avant tout de connaître les concepts et le vocabulaire de base, mais aussi de comprendre le sens, les valeurs et les objectifs de l'innovation engagée au regard des missions et activités permanentes et plus anciennes. Ce point est jugé essentiel pour favoriser l'appropriation et l'adhésion des acteurs du métier et éviter que l'équipe en première ligne sur le projet ne se trouve isolée ou marginalisée dans l'organisation de travail. C'est un levier essentiel de l'accompagnement au changement et de la modification des processus de travail qui seront induits sur le plus long terme²⁵¹. »

Il faut aussi prendre en compte le fait que l'ensemble de ces évolutions nécessite de s'appuyer à la fois sur des compétences techniques et informatiques et sur des compétences bibliographiques, plus traditionnelles mais néanmoins pointues. Ou encore « se pose également la question de gérer les tensions entre les agents chargés de la modélisation des données (travail plus conceptuel) et ceux chargés du traitement des données (travail plus opérationnel) ». Il peut être délicat de parvenir à faire dialoguer et travailler ensemble ces différents types de profils, ainsi que l'a montré Marc Scherer dans son mémoire intitulé « Bibliothécaires et informaticiens : convergences ou choc des

²⁵¹ BnF, *L'IBN à l'IFLA*, *op.cit.*

cultures ? »²⁵². Ces questions font écho à l'émergence de nouveaux métiers dans la profession et dont les contours sont encore à définir : *data manager* ou *data librarian*.

Animer des réseaux de managers

La réussite de la conduite du changement repose en grande partie sur la capacité des managers à susciter l'adhésion, à la fois celle des équipes mais aussi celle de la direction ou de la tutelle. Les managers doivent être capables de démontrer les bénéfices du changement pour l'organisation. Virgine Delaine souligne dans son mémoire sur l'accompagnement du changement en bibliothèque l'importance de proposer une vision. Les encadrants ont un rôle fondamental à jouer « dans la construction d'une vision incarnant l'évolution de leur bibliothèque sur le long terme »²⁵³. Si l'implication de la direction ou de la tutelle est essentielle pour porter le changement, elle doit être relayée par les managers de proximité dont le rôle consiste véritablement à faire de l'accompagnement au quotidien dans la vie des établissements.

Pour soutenir le personnel encadrant dans cette démarche, il pourrait être pertinent de mettre en place et d'animer un ou des réseaux de managers, à l'échelle nationale. Les managers de proximité doivent en effet être sensibilisés et informés afin de pouvoir à leur tour constituer des relais. Cela doit passer notamment par la fourniture d'outils tels que des supports d'information qu'ils pourront se réapproprier, l'essentiel étant que tous partagent le même niveau d'information afin de pouvoir porter les mêmes messages. Il s'agit aussi dans le cadre de tels réseaux de coordonner les actions de formation proposées aux agents.

Pour les mettre en œuvre, il existe déjà des réseaux sur lesquels s'appuyer, par exemple ceux des associations professionnelles : ADBU (Association des Directeurs de la Documentation et des Bibliothèques Universitaires), ADBGV (Association des directeurs des bibliothèques municipales et intercommunales des grandes villes de France), ADBDP (Association des Directeurs de Bibliothèques Départementales de Prêt). On peut également citer l'exemple de la BnF qui a mis en place en 2010 une organisation nommée ORHION (Observatoire des organisations et ressources humaines sous l'impact opérationnel du numérique), une instance de rencontre entre managers qui échangent sur l'évolution des métiers. Mentionnons également une tentative en ce sens mais plus informelle et à l'échelle internationale avec l'organisation dans le cadre d'une conférence satellite de l'IFLA²⁵⁴ d'un atelier pour les managers de bibliothèques visant à échanger sur des retours d'expériences²⁵⁵. Cet atelier avait l'ambition de commencer à constituer un réseau de managers qui pourrait au-delà de cette rencontre continuer à fournir un cadre pour ce type d'échanges.

Il serait sans doute intéressant de monter une structure de ce type à l'échelle nationale afin que tous les managers potentiellement impliqués puissent se

²⁵² SCHERER Marc, *Bibliothécaires et informaticiens : convergences ou choc des cultures ?*, mémoire DCB sous la direction de Christophe Evans, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/documents/64119-bibliothecaires-et-informaticiens-convergences-ou-choc-des-cultures.pdf>

²⁵³ DELAINE Virginie, *L'accompagnement du changement en bibliothèques : une approche managériale*, mémoire DCB sous la direction de Gilles Eboli, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/documents/64225-1-accompagnement-du-changement-en-bibliothèques-une-approche-managériale.pdf>

²⁵⁴ *Le web de données en bibliothèque, du projet à la pratique*, Conférence satellite IFLA 2014, Paris, BnF, 14 août 2014 <http://ifla2014-satdata.bnf.fr/program.html>

²⁵⁵ Proposé par Emmanuelle Bermès et Gildas Illien, le point de départ de cet atelier consistait en un retour d'expérience sur les projets qu'ils ont menés respectivement au Centre Pompidou avec le Centre Pompidou virtuel et à la BnF avec data.bnf.fr.

rencontrer à échéances régulières et le reste du temps disposer de canaux d'échanges dédiés. A notre avis c'est le chaînon indispensable à la bonne compréhension des changements à venir et des enjeux de la transition bibliographique.

CONCLUSION

Les actions entreprises par l'ABES et la BnF s'inscrivent dans une démarche qui va certainement dans la bonne direction. Leurs initiatives en faveur de l'ouverture et du rapprochement de leurs données permettent un décloisonnement qui fait qu'il est de moins en moins nécessaire à l'utilisateur de connaître les catalogues pour trouver des ressources, et qui rend donc celles-ci beaucoup plus visibles, ce qui est l'objectif que les bibliothèques ne doivent pas perdre de vue. Cette coopération s'inscrit dans un contexte particulièrement favorable : portée par les mouvements de fond de l'*open data* et du web de données elle peut bénéficier d'un calendrier propice avec les refontes des systèmes de production des deux agences. Il y a donc quelque chose comme un bon moment à saisir et à ne pas laisser passer. Car cette démarche semble aussi fragile. Elle concerne des sujets techniques requérant une certaine expertise, d'où les enjeux qui se posent en termes de vulgarisation et d'accompagnement du changement, mais aussi de formation de nouveaux experts. Par ailleurs cette démarche très récente n'est rendue possible que par la bonne volonté des uns et des autres. Certes, les objectifs liés à la coopération sont inscrits dans les projets d'établissements respectifs et dans une convention, mais il faut penser à la pérennisation de cette démarche, ainsi qu'à une réelle priorisation des actions. Les agences peuvent être confrontées à la difficulté de trouver des ressources pour instruire certains chantiers, compte tenu des enjeux auxquels elles doivent faire face, notamment le projet SGBM et le signalement des licences nationales pour l'ABES. Avec les problématiques qui leur sont propres elles peuvent aussi faire face à des divergences d'intérêt susceptibles de freiner la coopération. Pour donner un exemple la BnF pourrait être davantage attachée à la définition d'un profil français de RDA que l'ABES qui évolue dans un environnement plus international.

On peut reposer à ce stade la question de la mutualisation. Ouvrir ses données et les lier à d'autres ne doit pas dispenser d'envisager aussi la mutualisation de la production de ces données. L'existence de deux agences bibliographiques est problématique sur ce point et il est nécessaire de s'interroger sur les fonctions redondantes qui pourraient être mutualisées. En cela, la perspective d'un fichier national d'autorités qui permettrait à la BnF de dériver ses autorités du Sudoc est encourageante.

Les deux agences ont également un rôle fondamental à jouer dans la mutualisation à l'échelle nationale. La transition bibliographique en cours les positionne comme acteurs incontournables dans la fourniture de données de référence, en particulier d'autorité. Elles ont aussi vocation à partager les alignements de données réalisés : fournir des « bibliothèques d'alignements » et garantir leur validité. Elles doivent donc véritablement proposer une offre de réservoirs de données à disposition de toutes les bibliothèques, quel que soit leur type, ce qui serait un pas vers le décloisonnement.

Car il faut rappeler que ce paysage particulier comportant deux agences bibliographiques s'inscrit plus largement dans l'existence de deux réseaux, celui des bibliothèques publiques et celui des bibliothèques de l'enseignement supérieur et de la recherche, qui forment deux mondes professionnels bien distincts, reposant l'un sur la fonction publique d'État et l'autre sur la fonction publique territoriale.

L'annonce en mars 2014 de la fin du conventionnement entre le CNFPT et l'ENSSIB pour la formation des conservateurs territoriaux, signant la fin d'une formation commune aux deux corps, risque d'accentuer cette séparation. Il faudrait au contraire favoriser les rapprochements, et à ce titre, il est envisageable que la BnF apporte ses services à des bibliothèques de l'enseignement supérieur, ce qui entre tout à fait dans ses attributions de bibliothèque nationale, et que l'ABES à son tour puisse desservir les bibliothèques publiques notamment pour les métadonnées de ressources électroniques.

Faut-il souhaiter la fusion des deux agences bibliographiques ? S'il est indispensable de rechercher la mutualisation des fonctions, il ne faut pas oublier les spécificités de chacune et le fait qu'elles évoluent toutes deux dans un périmètre d'acteurs qui leur est propre. Et les besoins de mutualisation se posent aussi à l'intérieur de chacun de ces périmètres. C'est le cas dans le domaine de l'information scientifique et technique, particulièrement cloisonné, entre universités, grandes écoles, organismes de recherche. La coopération s'organise à ce niveau dans le cadre de la BSN, et il semblerait pertinent de questionner la possibilité d'un rapprochement de l'ABES avec l'INIST. Quant à la BnF elle évolue elle aussi au sein d'un réseau qui lui est propre, composé d'acteurs de la culture, de l'interprofession du livre ou encore des archives. Des pistes de mutualisation doivent aussi être recherchées dans ce cadre. Prenons l'exemple d'Électre, base de données professionnelle payante sur les livres disponibles à la vente ou à paraître et à laquelle les bibliothèques publiques sont obligées de recourir pour compléter les informations récupérées de la BnF ou parce que celles-ci ne sont pas disponibles assez tôt.

L'hypothèse hasardée au début de cette étude nous semble recevable : l'ABES et la BnF font des métiers différents et elles ont à ce titre une légitimité, celle de deux acteurs dont le rôle est appelé à devenir de plus en plus structurant, qui ont l'obligation d'accentuer leur collaboration et sur lesquels la communauté des bibliothèques va devoir compter de plus en plus.

Sources

ENTRETIENS

Pauline CHOUGNET (BnF) et Jérôme VILLEMINOZ (BnF), le 30 juin 2014

Françoise LERESCHE (BnF), le 8 octobre 2014

Olivier ROUSSEAU (ABES), le 15 octobre 2014

Philippe LE PAPE (ABES), le 15 octobre 2014

Yann NICOLAS (ABES) et Michael JEULIN (ABES), le 15 octobre 2014

Françoise BOURDON (BnF), le 23 octobre 2014

Marc MAISONNEUVE (Tosca Consultants), le 24 octobre 2014

Yves ALIX (IGB), le 24 novembre 2014

RÉUNIONS

Réunion opérationnelle ABES-BnF des 25 et 26 septembre 2014.

DOCUMENTS NON PUBLIÉS

Les documents ci-dessous ont été transmis par Gildas Illien :

- Convention de coopération entre l'Agence bibliographique de l'enseignement supérieur et la Bibliothèque nationale de France, signée le 24 juin 2014.
- Comptes-rendus des réunions opérationnelle ABES-BnF des 18 et 19 février 2014 et des 25 et 26 septembre 2014.
- BnF, *L'IBN à l'IFLA : Rapport de mission. Congrès de Lyon et Conférences satellites*, 2014.

JOURNÉES PROFESSIONNELLES

Open data et bibliothèques : une affaire de bibliothécaires ?, journée d'études organisée par les élèves promus de l'Enssib, Villeurbanne, 11 juin 2014.

Métadonnées et confiance : quelles normes pour les échanges entre communautés ?, Journée d'étude AFNOR CN46 / BnF, Paris, 27 juin 2014. Diaporamas et comptes rendus des communications disponibles sur http://www.bnf.fr/fr/professionnels/anx_journees_pro_2014/a.jp_2014_afnor.html

Le web de données en bibliothèque, du projet à la pratique, Conférence satellite IFLA 2014, Paris, 14 août 2014. Vidéos et textes des communications disponibles sur <http://ifla2014-satdata.bnf.fr/program.html>

Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?, session de la section Catalogage de l'IFLA, Lyon, 18 août 2014. Textes des communications disponibles sur <http://library.ifla.org/view/conferences/2014/2014-08-18/315.html>

Bibliographie

Tous les liens ont été vérifiés le 28 décembre 2014.

INSTITUTIONS

ABES, *Projet d'établissement 2012-2015*, novembre 2011. Disponible sur <http://www.abes.fr/Connaitre-l-ABES/Projet-d-etablissement>

AERES, Rapport d'évaluation de l'Agence bibliographique de l'enseignement supérieur, juillet 2012. Disponible sur : <http://www.aeres-evaluation.fr/Etablissements/AGENCE-BIBLIOGRAPHIQUE-DE-L-ENSEIGNEMENT-SUPERIEUR-ABES>

BÉRARD Raymond, « L' Abes à l'horizon 2016 », in *Bulletin des bibliothèques de France*, n° 1, janvier 2013. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2013-01-0072-016>

BnF, *Contrat de performance 2014-2016*, mai 2014. Disponible sur http://www.bnf.fr/fr/la_bnf/connaitre_bnf/a.contrat_performance.html

BRZUSTOWSKI Katie, BARRAL Sabine, « L'ABES, coordinatrice et animatrice de réseaux », in *Bulletin des bibliothèques de France*, n° 2, mars 2003. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2003-02-0032-007>

Comité d'histoire du ministère de la Culture, *Histoire administrative du Ministère de la culture et de la communication, 1959-2012. Les services de l'administration centrale*, Paris, La Documentation française, 2012.

Conseil d'État, *Rapport public 2012 du Conseil d'État - Les agences : une nouvelle gestion publique ?*, Paris, La Documentation française, 2012. Disponible sur <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/124000501/0000.pdf>

DUCHEMIN Pierre-Yves, avec la participation de Dominique LAHARY, *L'art d'informatiser une bibliothèque. Guide pratique*, Paris, Ed. du Cercle de la Librairie, 2000

ILLIEN Gildas, « 20 ans, le bel âge ! », in *Arabesques*, n°75, juillet 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-75>

MÉTIER

ABES, *Le blog de la formation*, [en ligne]. Montpellier, ABES. Disponible sur <http://laformation.abes.fr/>

AROT Dominique, CARBONE Pierre, CLAUD Joëlle, MICOL Charles, LECOQ Benoît, *Quels emplois dans les bibliothèques ? État des lieux et perspectives*, Inspection générale des bibliothèques, rapport n°2012-020, mars 2013. Disponible sur http://cache.media.enseignementsup-recherche.gouv.fr/file/2013/20/3/Quels_emplois_dans_les_bibliotheques_Etat_des_lieux_et_perspectives_247203.pdf

BnF, *Référentiel des emplois et des compétences*, 2012. Disponible sur http://www.bnf.fr/emploi/pdf/referentiel_emplois_compétences.pdf

DELAINÉ Virginie, *L'accompagnement du changement en bibliothèques : une approche managériale*, mémoire DCB sous la direction de Gilles Eboli, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/documents/64225-l-accompagnement-du-changement-en-bibliotheques-une-approche-managériale.pdf>

SCHERER Marc, *Bibliothécaires et informaticiens : convergences ou choc des cultures ?*, mémoire DCB sous la direction de Christophe Evans, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/documents/64119-bibliothécaires-et-informaticiens-convergences-ou-choc-des-cultures.pdf>

OPEN DATA

CARTIER Aurore, *Bibliothèques et Open data, Et si on ouvrait les bibliothèques sur l'avenir ?*, mémoire DCB sous la direction de Frédérique Schlosser, 2013. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/documents/60401-bibliotheque-et-open-data-et-si-on-ouvrait-les-bibliotheques-sur-l-avenir.pdf>

CHIGNARD Simon, *Open data : comprendre l'ouverture des données publiques*. Limoges, FYP éditions, 2012

ETALAB, *Vade-mecum sur l'ouverture et le partage des données publiques*, septembre 2013. Disponible sur <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/vademecum-ouverture.pdf>

GAILLARD Rémi, *De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de recherche ?*, mémoire DCB sous la direction de Monique Joly, 2014. Disponible sur <http://www.enssib.fr/bibliotheque-numerique/notices/64131-de-l-open-data-a-l-open-research-data-quelles-politiques-pour-les-donnees-de-recherche>

G8, *Charte du G8 pour l'Ouverture des Données Publiques*, 18 juin 2013. Disponible sur <http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/charte-g8-ouverture-donnees-publiques-fr.pdf>

Ministère de la Culture et de la Communication, *Ouverture et partage des données publiques culturelles. Pour une (r)évolution numérique dans le secteur culturel*, décembre 2013. Disponible sur <http://www.ladocumentationfrancaise.fr/rapports-publics/144000037/index.shtml>

Ministère de la Culture et de la Communication, *Guide Data Culture : Pour une stratégie numérique de diffusion et de réutilisation des données publiques numériques du secteur culturel*, mars 2013. Disponible sur <http://www.culture.fr/Professionnels/Reutilisation-de-ressources-Culture/Guide-Data-Culture>

TROJETTE Mohammed Adnène, *Rapport au Premier ministre. Ouverture des données publiques. Les exceptions au principe de gratuité sont-elles toutes légitimes ?*, juillet 2013. Disponible sur <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/134000739/0000.pdf>

SIGNALEMENT

ABES, BnF, *Rapport d'orientation pour le Comité stratégique bibliographique préparé par l'Agence bibliographique de l'enseignement supérieur et la Bibliothèque nationale de France*, novembre 2012. Disponible sur http://www.bnf.fr/documents/rapport_csb.pdf

ABES, BnF, *De RDA en France à la Transition bibliographique. Préconisations de l'ABES et de la BnF pour favoriser l'évolution des catalogues nationaux vers le web des données liées*, novembre 2014. Disponible sur http://www.bnf.fr/documents/1411_transition_bibliographique.pdf

ABES, *RDA@ABES*, [en ligne]. Montpellier, ABES. Disponible sur <http://rda.abes.fr/>

ABES, *Punktokomo ; blog technique de l'ABES*, [en ligne]. Montpellier, ABES. Disponible sur <http://punktokomo.abes.fr/>

ABES, *SGB mutualisé*, [en ligne]. Montpellier, ABES. Disponible sur <http://sgbm.abes.fr/>

ANDERSON Dorothy, *Universal Bibliographic Control. A long term policy. A plan for action*, PullachbeiMünchen, VerlagDokumentation, 1974

ANGJELI Anila, MAC EWAN Andrew, BOULET Vincent, « ISNI et VIAF transforment le paysage : pour des identités fiables et solides », in 80^e Congrès de l'IFLA, Session 86 - *Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?*, Lyon, 18 août 2014. Disponible sur <http://library.ifla.org/985/>

BnF. Pour les professionnels. Catalogage et indexation [en ligne]. Paris, Bibliothèque nationale de France. Disponible sur http://www.bnf.fr/fr/professionnels/normes_catalogage.html

BnF. Pour les professionnels. ISBD (International Standard Bibliographic Description) [en ligne]. Paris, Bibliothèque nationale de France. Disponible sur http://www.bnf.fr/fr/professionnels/normes_catalogage/a.normes_isbd_presentation.html

BnF. Pour les professionnels. RDA (Ressources: Description et Accès) [en ligne]. Paris, Bibliothèque nationale de France. Disponible sur http://www.bnf.fr/fr/professionnels/rda/s.rda_en_france.html?first_Art=non

BERMÈS Emmanuelle, « Bibliothèques, archives et musées : l'enjeu de la convergence des données du patrimoine culturel », in *Documentaliste-Sciences de l'information*, vol.48, n°4, décembre 2011, p.45-4

BERNON Jean, « Focus sur le projet SGBm », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canalc2.tv/video.asp?idvideo=12650>

BOBER Benjamin, « Focus sur BACON », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canalc2.tv/video.asp?idvideo=12652>

BOURDENET Philippe, « Enrichir le catalogue local avec des données exposées : interactions avec le catalogue national via web services », in 80^e

Congrès de l'IFLA, Session 86 - *Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?*, Lyon, 18 août 2014. Disponible sur <http://library.ifla.org/815/>

BOURDENET Philippe, « OCLC, l'histoire d'une coopération fructueuse », in *Documentaliste-Sciences de l'information*, vol.50, n°2, juillet 2013, p.30

BOURDON Françoise, ILLIEN Gildas, « A la recherche du temps perdu, retour vers le futur : CBU 2.0 », in 80^e Congrès de l'IFLA, Session 86 - *Universal Bibliographic Control in the Digital Age: Golden Opportunity or Paradise Lost?*, Lyon, 18 août 2014. Disponible sur <http://library.ifla.org/956/>

BOURDON Françoise, ILLIEN Gildas, ROCHE Mélanie, « Le catalogue des temps modernes, entre discipline et dissémination », in BARBIER Frédéric, DUBOIS Thierry, SORDET Yann, *De l'argile au nuage : une archéologie des catalogues*, catalogue d'exposition, Paris, Bibliothèque Mazarine (13 mars 2015 - 13 mai 2015), Paris, Bibliothèque Mazarine, Bibliothèque de Genève, Éditions des Cendres, 2015.

CARBONE Pierre, « Documentation numérique et réseaux documentaire », in *Arabesques*, n°74, avril 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-74>

CHOUGNET Pauline, « L'ISNI prend son essor », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

Comité d'orientation Rameau, Compte-rendu de la réunion du 2 décembre 2013, Paris, Bibliothèque nationale de France, janvier 2014. Disponible sur http://rameau.bnf.fr/informations/pdf/comite_orientation_rameau_cr_20131202.pdf

Fédération Internationale des Associations de Bibliothécaires et des Bibliothèques, *ISBD : Description bibliographique internationale normalisée*, édition intégrée, Paris, Bibliothèque nationale de France, 2013. Disponible sur http://www.bnf.fr/documents/isbd_trad_francais.pdf

FONS Ted, « Interview with Gildas Illien », in *Information Standards Quarterly*, Vol.25, n°4, Hiver 2013. Disponible sur <http://www.niso.org/publications/isq/2013/v25no4/illien/>

GIRARD Christine, LECOQ Benoît, *L'avenir des catalogues collectifs*, Inspection générale des bibliothèques, rapport n°2011-017, juillet 2011. Disponible sur : http://cache.media.enseignementsup-recherche.gouv.fr/file/2011/86/1/Rapport.cat-coll.-V-25-11-2011_200861.pdf

Groupe de préfiguration projet Sudoc II, « Sudoc II. Année 0 », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

Groupe de travail IFLA sur les Fonctionnalités requises des notices bibliographiques, *Fonctionnalités requises des notices bibliographiques : rapport final*, Paris, Bibliothèque nationale de France, 2012. Disponible sur http://www.bnf.fr/documents/frbr_rapport_final.pdf

Groupe de travail IFLA sur les Fonctionnalités requises et la numérotation des notices d'autorité (FRANAR), *Fonctionnalités requises des données d'autorité : un modèle conceptuel : rapport final*, Paris, Bibliothèque nationale de France, 2010. Disponible sur http://www.bnf.fr/documents/frad_rapport_final.pdf

Groupe de travail IFLA sur les Fonctionnalités requises des notices d'autorité matière (FRSAR), *Fonctionnalités requises des données d'autorité matière : un modèle*

conceptuel, Paris, Bibliothèque nationale de France, 2012. Disponible sur http://www.bnf.fr/documents/frsad_rapport_final.pdf

Groupe technique sur l'adoption de RDA en France, *Rapport du groupe technique sur l'adoption de RDA en France*, Rapport au Groupe stratégique pour sa réunion du 25 mars 2011. Disponible sur http://rda-en-france.enssib.fr/sites/rda-en-france.enssib.fr/files/Rapport%20au%20groupe%20strat%C3%A9gique%20pour%20r%C3%A9union%2025%20mars_v3_20110323_0.pdf

IFLA Working Group on GARE Revision, *Guidelines for Authority Records and References*, München, K.G. Saur, 2001. Disponible sur <http://www.ifla.org/files/assets/hq/publications/series/23.pdf>

ILLIEN Gildas, « Décrire les objets du savoir, les nouveaux paradigmes du catalogage » in *Documentaliste-Sciences de l'Information*, vol.50, n°3, octobre 2013, p.26-29

ILLIEN Gildas, « Metadata mining : fouiller les données des catalogues », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

JSC, *Resource description & access RDA / developed in a collaborative process led by the Joint Steering Committee for Development of RDA (JSC)*, Chicago, American Library Association, 2010

KALFON Jérôme, « La place des réseaux de l'Abes à l'heure du décloisonnement », in *Arabesques*, n°74, avril 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-74>

LEMELLE Françoise. « Le serveur bibliographique national », in *Bulletin des bibliothèques de France*, n° 3, mai 1991. Disponible sur <http://bbf.enssib.fr/consulter/bbf-1991-03-0224-002>

LE PAPE Philippe, avec la collaboration de Françoise LERESCHE, « Le signalement des documents. Catalogues informatisés et communication de données », in SANZ Pascal, *Guide de la coopération entre bibliothèques*. Paris, Éd. du Cercle de la librairie, 2008.

LE PAPE Philippe, « Sudoc, un instant de perplexité », [enregistrement vidéo], in Journées ABES 2013, Montpellier, 15 mai 2013. Disponible sur <http://canalc2.u-strasbg.fr/video.asp?idvideo=11836>

LERESCHE Françoise, *Avons-nous besoin de RDA pour mettre les catalogues de bibliothèques sur le web de données ? : la réflexion française sur l'évolution des catalogues et des règles de catalogage*, 2012. Disponible sur <http://www.bn.gov.ar/descargas/catalogadores/encuentro2011/ponencia-23-A-Leresche.pdf>

LERESCHE Françoise, « Qu'est-ce que le RDA ? », [enregistrement vidéo], in 5 à 7 de l'ADBS, *RDA : le nouveau paradigme du catalogage ?*, Paris, 29 mars 2011. Disponible sur <http://www.adbs.fr/rda-le-nouveau-paradigme-du-catalogage-2-qu-est-ce-que-le-rda--102462.htm>

LERESCHE Françoise, « Réunion annuelle d'EURIG », in *Bulletin des bibliothèques de France*, n° 6, novembre 2013. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2013-06-0077-004>

LESCURE Pierre, *Mission « Acte II de l'exception culturelle » : contribution aux politiques culturelles à l'ère numérique*, Ministère de la Culture et de la

Communication, mai 2013. Disponible sur http://www.culturecommunication.gouv.fr/var/culture/storage/culture_mag/rapport_lescur/index.htm

LIARD Émilie, « RDA vu de l'intérieur », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 21 mai 2014. Disponible sur <http://canalc2.u-strasbg.fr/video.asp?idvideo=12655>

Ministère de la Culture et de la Communication, Boîte à outils du numérique en bibliothèque, Fiche pratique n°1, La récupération des données pour le catalogue, 2012. Disponible sur <http://www.culturecommunication.gouv.fr/Disciplines-secteurs/Livre-et-Lecture/Bibliotheques/Numerique-et-bibliotheques/Boite-a-outils-du-numerique-en-bibliotheque>

MISTRAL François, « Articulation ISNI-IdRef : un enjeu pour l'identification pérenne », in *Arabesques*, n°76, octobre 2014. Disponible sur <http://www.abes.fr/Arabesques/Arabesques-n-76>

MKADMI Abderrazak, SALEH Imad, *Bibliothèque numérique et recherche d'informations*, Paris, Hermès Science, Lavoisier, 2008

NICOLAS Yann, « L'Abes engage ses forces et ses données », in *Documentaliste-Sciences de l'information*, vol.48, n°4, décembre 2011, p.51

NICOLAS Yann, « Calames, et après ? », in *Bulletin des bibliothèques de France*, n° 6, novembre 2008. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2008-06-0029-005>

NICOLAS Yann, « Focus sur le hub de métadonnées », [enregistrement vidéo], in Journées ABES 2014, Montpellier, 20 mai 2014. Disponible sur <http://www.canalc2.tv/video.asp?idvideo=12651>

NISO, *Understanding Metadata*, 2004. Disponible sur <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

SIMON Agnès, WENZ Romain, « Des outils automatiques pour le signalement en bibliothèque : expérimentations autour du projet data.bnf.fr ». in *Bulletin des bibliothèques de France*, n°5, septembre 2012. Disponible sur <http://bbf.enssib.fr/consulter/bbf-2012-05-0039-008>

WENZ Romain, « Data.bnf.fr : la fabrique », in *Synthèse de la journée d'étude du GFII : « Données culturelles et Linked Open Data : valoriser le patrimoine public dans le web des données »*, 26 mars 2013, Maison de l'Europe. Disponible sur <http://www.gfii.fr/fr/document/journee-d-etude-du-gfii-donnees-culturelles-et-linked-open-data-valoriser-le-patrimoine-public-dans-le-web-de-donnees>

WILLER Mirna, DUNSIRE Gordon, *Bibliographic information organization in the semantic web*, Oxford : Chandos Publishing, 2013.

WEB, WEB SÉMANTIQUE, BIBLIOTHÈQUES ET WEB DE DONNÉES

ANGJELI Anila, « La normalisation en bibliothèque à l'heure du web sémantique » in *Documentaliste-Sciences de l'information*, vol.48, n°4, décembre 2011, p.53

BERMÈS Emmanuelle, avec la collaboration d'Antoine ISAAC et Gautier POUPEAU, *Le Web sémantique en bibliothèque*, Paris, Ed. du Cercle de la Librairie, 2013.

BERMÈS Emmanuelle, *Figoblog*, [en ligne]. Disponible sur <http://www.figoblog.org/>

BERNERS-LEE Tim, *Linked Data*, 2006. Disponible sur <http://www.w3.org/DesignIssues/LinkedData.html>

BERNERS-LEE Tim, « Le web à venir », [enregistrement vidéo], in TED Conférences, février 2009. Disponible sur http://www.ted.com/talks/tim_berniers_lee_on_the_next_web?language=fr#

DUHAMEL Benjamin, *Les technologies du web sémantique et du record linkage au service de data.bnf.fr et du Linked Open Data culturel : Étude sur les nouveaux paradigmes informationnels*, mémoire de Master à l'Université Lille 3 sous la direction de Gildas Illien et Gérard Kembellec, 2014.

Groupe d'incubation W3C, *Rapport final du groupe d'incubation "Bibliothèques et web de données"*, rapport du 25 octobre 2011. Disponible en français sur <http://mediatheque.cite-musique.fr/MediaComposite/ARTICLES/W3C/XGR-ld-fr.html>

Table des matières

SIGLES ET ABRÉVIATIONS.....	7
INTRODUCTION.....	9
PRÉAMBULE : CONTEXTE ET DÉFINITIONS.....	11
Qu'est-ce que l'open data ?.....	11
Le contexte français.....	12
Les métadonnées de bibliothèques.....	15
Ouverture juridique.....	15
Ouverture technique.....	16
LES INSTITUTIONS EN PRÉSENCE.....	18
Deux agences bibliographiques.....	18
La Bibliothèque nationale de France.....	19
<i>Missions.....</i>	<i>20</i>
<i>Le réseau de la BnF.....</i>	<i>20</i>
L'échec du catalogue partagé.....	20
Pôles associés et projets de coopération documentaire.....	21
La mise à disposition de produits bibliographiques.....	22
<i>L'agence bibliographique nationale.....</i>	<i>23</i>
Le contrôle bibliographique universel.....	23
Du CBN au département de l'IBN.....	23
<i>Actualité de la BnF.....</i>	<i>24</i>
Faire évoluer le dépôt légal.....	24
Dérivation des notices d'OCLC.....	25
Une nouvelle interface pour le catalogue.....	26
Le projet ReLire.....	26
Vers la transition bibliographique.....	26
L'Agence Bibliographique de l'Enseignement Supérieur.....	27
<i>La naissance de l'ABES.....</i>	<i>27</i>
Création de l'agence.....	27
Le Sudoc.....	27
L'animation du réseau.....	28
Bilan.....	29
<i>La diversification des activités de l'ABES.....</i>	<i>29</i>
Calames.....	30
La stratégie des API.....	30
La valorisation des thèses.....	30
L'acquisition de documentation électronique.....	31
De nouvelles missions pour l'ABES.....	31
Un projet d'établissement.....	32
<i>Les nouveaux projets.....</i>	<i>32</i>
Bacon.....	32
Le hub de métadonnées.....	33
SGBm et Sudoc II.....	33
Les débuts de la coopération.....	35
<i>Le signalement des ressources.....</i>	<i>35</i>
Les échanges de notices.....	35
Élaboration de Rameau.....	35
Le Catalogue Collectif de France.....	36

<i>La mise en commun d'une réflexion stratégique</i>	36
<i>Des rôles bien définis</i>	37
LA TRANSITION BIBLIOGRAPHIQUE	39
L'évolution des collections, des besoins et des usages	39
<i>Les ressources numériques changent la donne</i>	39
<i>Le web change la donne</i>	40
<i>Bibliothèques et web de données</i>	42
Vers un nouveau modèle pour la description bibliographique	45
<i>Nos outils actuels</i>	45
Les règles de description bibliographique.....	45
Les formats d'encodage.....	46
<i>Une nouvelle façon de modéliser l'information : FRBR</i>	46
Qu'est-ce que le modèle FRBR ?.....	46
Quels sont les avantages des FRBR ?.....	48
Le rôle central des autorités.....	48
<i>Un nouveau code de catalogage : RDA</i>	49
Qu'est ce que RDA ?.....	49
Adopter RDA ?.....	50
<i>Vers une révision des formats MARC</i>	51
Vers de nouveaux modes de travail ?	53
<i>Du catalogage au « cataliage »</i>	53
<i>Une nouvelle répartition du travail entre les acteurs</i>	56
OUVRIR LES DONNÉES : LES STRATÉGIES MISES EN ŒUVRE	59
S'insérer dans le web de données	60
<i>L'approche de l'ABES</i>	61
Principes et éléments techniques.....	61
Quelles réutilisations ?.....	62
<i>L'approche de la BnF</i>	63
Principes et éléments techniques.....	63
Quelles réutilisations ?.....	65
<i>Bilan</i>	65
<i>Interrogation croisée des données</i>	66
Mutualiser les autorités	66
<i>Les personnes et collectivités</i>	67
Contexte : des référentiels internationaux.....	67
VIAF.....	67
ISNI.....	68
ORCID.....	69
Alignement des identifiants ABES et BnF.....	70
<i>Les sujets</i>	71
<i>Les titres et œuvres</i>	72
<i>Vers un fichier national d'autorités ?</i>	74
Préparer la transition bibliographique	74
<i>Un nouveau cadre pour adopter RDA</i>	74
<i>Conduire le changement</i>	76
Caractériser le changement.....	77
Informer, communiquer.....	78
Former, gérer les compétences.....	79
Animer des réseaux de managers.....	80
CONCLUSION	82

SOURCES	85
BIBLIOGRAPHIE	86
Institutions	86
Métier	86
Open data	87
Signalement	88
Web, web sémantique, bibliothèques et web de données	91
TABLE DES MATIÈRES	93