

Remerciements

Je souhaite vivement remercier Catherine Lupovici, directrice du département de la bibliothèque numérique, pour m'avoir accueilli dans son département, au sein duquel j'ai pu découvrir toute la richesse des réflexions sur le document numérique.

Je suis tout particulièrement redevable à Emmanuelle Bermès, ainsi qu'à Frédéric Martin, tous deux conservateurs au département de la bibliothèque numérique, pour m'avoir dirigé tout au long de ce stage et l'avoir soigneusement préparé. Ils ont toujours été disponibles pour me faire partager leurs connaissances.

J'adresse aussi ma gratitude aux agents du département (Sara Aubry, Thierry Cloarec, Gildas Illien, France Lasfargues, Noémie Lesquins), au milieu desquels il a été très agréable de travailler, pour leur accueil chaleureux et leur enthousiasme.

Je suis très reconnaissant envers Valérie Game et Sophie Sepetjan, chargées du service juridique de la BnF, pour leur soutien, et pour le temps qu'elles ont consacré au travail sur les droits des documents numériques.

Enfin, je tiens à remercier tous les membres de la BnF que j'ai eu l'occasion de rencontrer au cours de mon stage, particulièrement les agents des départements de la coopération et de la conservation, pour m'avoir aidé à découvrir leur établissement et leur métier.

Sommaire

INTRODUCTION.....	6
PARTIE 1 : LE DÉPARTEMENT DE LA BIBLIOTHÈQUE NUMÉRIQUE... 7	
1. HISTORIQUE.....	7
1.1. <i>Origine</i>	7
1.2. <i>La réorganisation</i>	8
2. LES PÔLES.....	8
2.1. <i>Le pôle « bibliothéconomie numérique »</i>	9
2.1.1. <i>La réflexion sur les accès et les services</i>	9
2.1.2. <i>L'OAI à la BnF</i>	9
2.1.2.1. <i>Fédérer les ressources</i>	9
2.1.2.2. <i>Faciliter les partenariats thématiques</i>	10
2.1.2.3. <i>Être plus visible sur le Web</i>	10
2.1.3. <i>Un travail de normalisation</i>	11
2.1.4. <i>La réflexion sur la bibliothèque numérique européenne</i>	12
2.1.5. <i>Les enjeux d'une évolution</i>	12
2.2. <i>Le pôle « DL Web »</i>	13
2.2.1. <i>Enjeux et défis</i>	13
2.2.2. <i>Les premières expérimentations</i>	14
2.2.3. <i>Le projet de loi relatif au dépôt légal</i>	15
2.2.4. <i>Collecter</i>	15
2.2.4.1. <i>Sélection de sites</i>	15
2.2.4.2. <i>Dépôt</i>	16
2.2.4.3. <i>Capture automatisée par robot</i>	16
2.2.5. <i>Traiter</i>	17
2.2.5.1. <i>Contrôle qualité</i>	17
2.2.5.2. <i>La conservation</i>	17
2.2.6. <i>L'accès</i>	18
3. TRAVAILLER AU DBN.....	18
3.1. <i>Moyens humains et financiers</i>	18
3.2. <i>Le travail au quotidien</i>	19

<u>PARTIE 2 : LA CHAÎNE DE TRAITEMENT DES DOCUMENTS</u>	
<u>NUMÉRIQUES À LA BNF.....</u>	<u>22</u>
1. GALICA.....	22
<i>1.1. Sélection et acquisition.....</i>	<i>23</i>
1.1.1. Critères de sélection.....	23
1.1.2. Origine des ouvrages.....	24
1.1.3. Budget.....	24
<i>1.2. Numérisation.....</i>	<i>25</i>
1.2.1. Les filières de numérisation.....	25
1.2.2. Vérification et catalogage.....	25
1.2.3. Contrôle qualité.....	25
<i>1.3. Migration et consultation.....</i>	<i>26</i>
1.3.1. La migration sur les serveurs du DSI.....	26
1.3.2. L'interface de consultation.....	26
2. LES BASES DU DÉPARTEMENT DE LA REPRODUCTION.....	27
<i>2.1. Alimentation de la base.....</i>	<i>27</i>
<i>2.2. Description et mise en ligne.....</i>	<i>27</i>
3. L'ARCHIVAGE PÉRENNE ET LE MAGASIN NUMÉRIQUE.....	28
<i>3.1. La variété des producteurs de données numériques.....</i>	<i>28</i>
<i>3.2. Difficultés rencontrées.....</i>	<i>29</i>
3.2.1. Problèmes de supports et de formats.....	29
3.2.2. Métadonnées.....	30
<i>3.3. Modèle cible.....</i>	<i>30</i>
<u>PARTIE 3 : ACTIVITÉS EFFECTUÉES PAR LE STAGIAIRE.....</u>	<u>32</u>
1. LA RÉFLEXION SUR LE STATUT JURIDIQUE DU DOCUMENT NUMÉRIQUE.....	32
<i>1.1. La législation confrontée au défi du numérique et de l'Internet.....</i>	<i>32</i>
1.1.1. La législation en vigueur.....	32
1.1.2. Le défi présenté par les nouvelles technologies.....	33
1.1.3. La réaction : renforcement de l'arsenal juridique et technique.....	34
<i>1.2. L'influence de ce nouveau contexte à la BnF.....</i>	<i>34</i>
<i>1.3. Travail effectué par le stagiaire.....</i>	<i>35</i>

1.3.1. Les arbres de décision.....	35
1.3.2. L'expression des droits des ressources numériques mises en ligne par la BnF.....	36
1.3.3. Apport de ce travail.....	37
2. AUTRES ACTIVITÉS.....	37
2.1. <i>Participation aux activités du service.....</i>	<i>37</i>
2.2. <i>Découverte de la BnF.....</i>	<i>38</i>
CONCLUSION.....	40
BIBLIOGRAPHIE.....	41
TABLE DES ANNEXES.....	44

Introduction

La Bibliothèque nationale de France m'a accueilli au sein du département de la bibliothèque numérique, du 5 septembre au 26 novembre 2005, dans le cadre d'un stage d'étude de l'Enssib. J'y ai travaillé sous la responsabilité d'Emmanuelle Bermès, conservatrice dans le pôle « bibliothéconomie numérique ».

Ce stage présentait de multiples intérêts. Je souhaitais découvrir les problématiques de la numérisation, point de jonction entre deux domaines qui m'attirent : le patrimoine et l'informatique documentaire. De plus, c'était l'occasion de mieux connaître la BnF, établissement aux missions nationales et internationales, dont l'expérience en matière de numérisation était reconnue. D'autre part, on m'avait proposé de travailler sur les aspects juridiques de la numérisation, c'est-à-dire d'étudier les droits régissant la communication des documents numériques, et la façon de les exprimer. J'étais intéressé par le caractère assez transversal de ce sujet, puisque les problèmes juridiques prennent une part croissante dans la question de la collecte, de la conservation et de l'accès aux documents numériques mais aussi analogiques. De plus, un tel sujet trouve de nombreux échos dans l'actualité, qu'il s'agisse de la polémique autour de Google Print et ses rebondissements judiciaires ou de la future transposition, dans la loi française, de la directive européenne sur les droits d'auteurs. Toutefois, comme j'ai mené de façon concomitante des recherches dans le cadre d'une thèse de doctorat, je n'ai pu consacrer qu'un mi-temps à mon stage.

Le rapport de stage se découpe en trois parties. La première présente mon département d'accueil (histoire, fonctionnement moyens). La seconde partie concerne la chaîne de traitement des documents numériques, particulièrement ceux de Gallica. Elle adopte ainsi une approche plus verticale, en abordant l'activité de plusieurs départements (bibliothèque numérique, coopération, conservation, systèmes d'information). Enfin, la troisième partie expose brièvement mon travail au cours de ce stage, en présentant les influences du nouveau contexte juridique et technique sur la numérisation, la réalisation du travail demandé par le département, et enfin les autres activités auxquelles j'ai été employé.

Partie 1 : le département de la bibliothèque numérique

Le département de la bibliothèque numérique (DBN) fait partie de la Direction des services et des réseaux (DSR) créée à la fin de l'année 1998. Cette direction regroupe plusieurs autres services transversaux comme l'Agence bibliographique nationale (ABN), le département du dépôt légal, le département de la coopération, le département de la conservation ou encore le département des systèmes d'information (DSI). Elle est, avec la Direction des collections (DCO) et la Direction de l'administration et du personnel, l'une des trois directions de la BnF.

1. Historique

1.1. Origine

Le département est de création relativement récente. Décidé en 1998, il a commencé à fonctionner à partir du premier trimestre de 1999.

Le choix d'un tel département était dès l'origine d'ordre stratégique et politique. Il correspondait à la réalisation d'un des objectifs initiaux de la Bibliothèque de France, « bibliothèque d'un genre entièrement nouveau »¹. En effet, dans l'esprit de ses concepteurs, la nouvelle bibliothèque, alliant modernité technologique et volonté de démocratiser la recherche, devait produire des documents numériques, qui permettaient de ne plus manipuler l'ouvrage original, et qui pouvaient être diffusés instantanément à n'importe quel usager. En outre, la création de ce département marquait la volonté forte de la BnF de s'inscrire au sein du réseau Internet, alors en pleine expansion.

Initialement, le département de la bibliothèque numérique était chargé de toutes les étapes de production et de diffusion des documents numériques, d'un bout à l'autre de la chaîne.

¹ François STASSE. *La véritable histoire de la grande bibliothèque*. Paris : Éd. du Seuil, 2002, p. 26.

1.2. La réorganisation.

Cette organisation « verticale », qui considérait le document numérique comme un document spécifique, nécessitant une structure adéquate, a fait place, à la suite d'une réorganisation entamée en 2003, à un mode de fonctionnement plus horizontal. Les différentes tâches de traitement du document ont été attribuées à des services qui disposaient de compétences similaires :

- La sélection des documents a désormais dépendu du département de la coopération. En effet, une des vocations des « pôles associés », dont ce département a la charge, est de partager la tâche de conservation du patrimoine français, notamment par la numérisation en partenariat².
- Les aspects techniques (contrôle qualité, gestion des prestataires...), sont passés au département de la conservation.
- Le DSI a été chargé de la production de documents.

A la suite de cette réorganisation, il a été nécessaire de repenser les objectifs et les fonctions du DBN. Il fut décidé d'en faire en quelque sorte un département « pilote », dont les agents devaient consacrer leur activité à des travaux de veille, d'expérimentation et de réflexion. Ainsi, un rôle fort d'expertise et d'innovation devait incomber au DBN.

2. Les pôles

En raison de sa taille réduite, le département de la bibliothèque numérique n'a pas été subdivisé en services. En revanche, il a été jugé nécessaire de distinguer deux « pôles », correspondant aux deux vocations du DBN. D'une part, le pôle « bibliothéconomie numérique », principalement chargé de conduire des projets ayant trait aux procédures d'accès aux documents numérique et à la définition des métadonnées nécessaires. D'autre part, le pôle « dépôt légal du Web », destiné à expérimenter et mettre en place les procédures de collecte, de conservation et d'accès du Web « français ». Nés d'un remaniement récent, appelés à évoluer

² Voir *infra*, p. 23.

fréquemment en raison des rapides changements dans le domaine technique, ces pôles étaient en pleine redéfinition lorsque j'ai effectué mon stage. J'ai ainsi eu l'occasion de constater quelles interrogations pouvaient présider à la réorganisation d'un département.

2.1. Le pôle « bibliothéconomie numérique »

2.1.1. La réflexion sur les accès et les services

Les accès et les services représentent les aspects « fonctionnels » du document numérique. Ainsi, c'est le DBN, en collaboration avec le département de la coopération, qui traite des questions relatives à l'interface de Gallica. Les usages des lecteurs doivent être pris en compte, tout comme les expériences menées par des institutions similaires.

2.1.2. L'OAI à la BnF

Les travaux menés autour du protocole OAI-PMH correspondent eux aussi à une réflexion sur un meilleur accès aux documents numériques de la BnF³. L'entrepôt OAI fonctionne de façon expérimentale depuis 2004. Il sera mis en production officielle courant 2006. Il sera alors à même de répondre à divers objectifs.

2.1.2.1. Fédérer les ressources

Le premier est de permettre à un utilisateur, à partir du site de la BnF, de lancer des requêtes interrogeant l'ensemble des ressources numériques de la bibliothèque. En effet, de multiples sources de documents numériques sont disponibles sur le site de la BnF :

- Gallica.
- Mandragore, base iconographique du département des manuscrits.
- La banque d'images, base de données de consultation des images numérisées par le département de la reproduction.
- Les Expositions virtuelles.

³ Pour une présentation du fonctionnement du protocole OAI-PMH, voir **François NAWROCKI**. *Le protocole OAI et ses usages en bibliothèque* [en ligne]. Paris : Ministère de la culture et de la communication, 2005. Mise à jour le 15 février 2005. [Consulté le 23 novembre 2005]. Disponible sur : <<http://www.culture.gouv.fr/culture/dll/OAI-PMH.htm>>.

- Les Dossiers pédagogiques.
- Sans oublier l'ensemble de la documentation professionnelle contenue sur le site de la BnF, qui n'est pas encore prise en compte au sein des ressources accessibles dans l'entrepôt OAI.

Ainsi, la BnF se trouverait dans le rôle d'un fournisseur de service, mais uniquement pour ses propres ressources, selon le modèle de la bibliothèque nationale d'Australie.

2.1.2.2. Faciliter les partenariats thématiques

Protocole d'échange de métadonnées, le protocole OAI permet de réunir aisément, sans lourds transferts de fichiers, des ressources hébergées par les serveurs de différentes institutions. Le dossier France-Amérique, mis en ligne en 2005, est une des plus belles réalisations. L'objectif était de monter un site thématique, entre la BnF et la Bibliothèque du Congrès, au sujet de la présence française en Amérique. Chaque établissement s'est chargé, de son côté, de proposer des documents et de les numériser. Une interface utilisant le protocole OAI permet à l'utilisateur, à partir du site Gallica ou du site de la Bibliothèque du Congrès, de faire une recherche sur les ressources des deux établissements.

2.1.2.3. Être plus visible sur le Web

Actuellement, les ressources disponibles dans Gallica, aussi bien que les notices du catalogue BN-Opale Plus, relèvent du « Web profond », qui n'est pas indexé par les moteurs de recherche.

Néanmoins, ces moteurs, à l'instar de M.S.N., Yahoo et Google, cherchent à obtenir, grâce au protocole OAI, des références de documents indexées en Dublin Core. Ceci représente un moyen pour ces moteurs de proposer à leurs usagers des ressources de qualité – les références provenant d'un entrepôt OAI sont souvent bien classées dans les pages des moteurs de recherche, puisque les institutions qui mettent en place ces entrepôts présentent des gages de sérieux.

Pour l'instant, le référencement des ressources numériques de la BnF grâce à l'OAI n'est encore que dans sa phase préparatoire : sur 1,5 millions de documents

visualisés sur Gallica tous les mois, 57 000 seulement l'ont été via l'OAI. Toutefois, ces chiffres sont appelés à croître.

De surcroît, la totalité des 12 millions de notices bibliographiques de BN-Opale Plus devrait, à terme, être disponible dans un autre entrepôt OAI. A cette fin, il va falloir donner à chaque notice en InterMarc un équivalent en Dublin Core. Au regard du nombre de notices, la seule solution envisageable est celle de grands chantiers de conversion automatique (*mapping*) de l'InterMarc en Dublin Core. Les règles de *mapping* définies par les membres du DBN doivent indiquer quel champ de l'InterMarc doit être transféré dans tel champ du DC. La difficulté est qu'il faut prendre en compte toutes les stratifications du catalogage : des notices issues d'une conversion rétrospective ne vont pas être parfaitement identiques à des notices créées en 2005. On doit prévoir des règles de *mapping* pour chaque cas de figure.

2.1.3. Un travail de normalisation

Cet axe de travail contribue à donner une partie de son sens au pôle document numérique, et justifie la dimension d'expertise du département. Le DBN a pour vocation de formaliser et d'éditer des documents de référence. Pour éditer un guide de « bonnes pratiques » pour la publication des documents sur le site bnf.fr, par exemple, il est en premier lieu nécessaire d'analyser les usages recensés sur le site. Ensuite, il faut faire un bilan de ces usages, et dégager les meilleures procédures. Cela aboutit à donner des conseils sur les polices à adopter, l'encodage, les formats... Autre exemple, des agents du DBN ont participé à la traduction de la norme OAIS en français⁴.

Comme cette activité de normalisation a une vocation aussi bien interne qu'externe, le DBN doit aussi assurer la diffusion de ces recommandations. A cette fin, il s'occupe d'un site, Bibnum⁵, qui va être bientôt refondu.

⁴ Voir *infra*, p. 30.

⁵ **BIBLIOTHEQUE NATIONALE DE FRANCE. *Bibnum***. [en ligne]. Paris, BnF, 2004. Mise à jour octobre 2005. [Consulté le 9 décembre 2005]. Disponible sur : <<http://bibnum.bnf.fr>>

2.1.4. La réflexion sur la bibliothèque numérique européenne

En janvier 2005, Jean-Noël Jeanneney, président de la BnF, réagissait vivement à l'annonce du projet Google Print, qui visait à numériser 15 millions d'ouvrages de six grandes bibliothèques anglo-saxonnes. Il dénonçait « le risque d'une domination écrasante de l'Amérique dans la définition que les prochaines générations se feront du monde »⁶. Il en appelait à une réaction politique, à l'échelle européenne.

C'est pourquoi, le 13 juillet 2005, était mis en place par Renaud Donnedieu de Vabres, ministre de la culture et de la communication, le comité de pilotage du projet de Bibliothèque numérique européenne (BNE), organe interministériel souhaité par le président de la République, chargé, au niveau national, de « *l'examen de la faisabilité et des modalités de mise en œuvre du projet de bibliothèque numérique européenne* », selon le communiqué officiel.

Un agent du DBN a été affecté à temps plein auprès de la secrétaire générale de ce comité de pilotage, Valérie Tesnière, directrice du département de la coopération.

Au sein du comité, cinq groupes de travail ont été mis en place :

- état des lieux des bibliothèques numériques
- contenu et politique documentaire de la future BNE
- publics cibles
- architecture du site
- modèles de financement

Le travail du secrétariat général est de coordonner les travaux, de fixer les ordres du jour des réunions des groupes de travail, et de formaliser leurs décisions. Les principales conclusions seront connues à la fin du mois de décembre, avec la remise d'un rapport final au président de la république, Jacques Chirac.

2.1.5. Les enjeux d'une évolution

Le pôle « bibliothéconomie numérique », anciennement baptisé « document numérique », a souhaité ce changement de nom pour mieux définir ses fonctions. Il s'agit de démontrer la cohérence des activités du pôle : intégrer le document

⁶ Jean-Noël JEANNENEY. « Quand Google défie l'Europe ». *Le Monde*, 23 janvier 2005.

numérique, comme son équivalent analogique, dans un ensemble global couvrant aussi bien l'archivage que la gestion des entrées, ou l'accès. Or, bien souvent, en terme de numérique, seule la question de la conservation est actuellement abordée. Une même réflexion a amené des bibliothécaires anglo-saxons à proposer le concept de « *digital curation* » en remplacement de « *digital preservation* ».

L'idée de créer un pôle de bibliothéconomie numérique au DBN revient à proposer au sein de l'établissement un pôle visible pour l'expertise et la coopération sur des projets du domaine numérique à composante bibliothéconomique. Ce pôle pourrait ainsi être amené à piloter la mise en place d'un vaste « magasin numérique », c'est-à-dire un système qui devra assurer le stockage, la gestion et la communication de l'ensemble des documents numériques de la BnF⁷. Le pôle pourra notamment être chargé de favoriser l'intégration à ce magasin de la collection gérée par l'autre pôle, les sites Web acquis au titre du futur dépôt légal.

2.2. Le pôle « DL Web »

2.2.1. Enjeux et défis

Conserver la trace de ce qui circule sur Internet est devenu un enjeu patrimonial fondamental. Si le « réseau des réseaux » permet de diffuser aisément, à faible coût, et dans le monde entier, une masse d'informations gigantesque et en perpétuel renouvellement, il manque singulièrement de mémoire. Ce problème est d'autant plus important que de plus en plus, la documentation qui était traditionnellement disponible sur support analogique tend à être substituée par une documentation en ligne. Mais cette conservation représente également un défi majeur. La masse des documents à traiter, la difficulté de les rechercher sur la Toile, la nécessité de les stocker à long terme, malgré l'évolution des techniques, des formats et des supports, et celle de les communiquer aux citoyens, représentent un défi majeur.

C'est au titre du dépôt légal que la Bibliothèque nationale de France se devait de relever ce défi. En effet, initialement promulgué pour les livres imprimés, le dépôt légal s'est, au cours de l'Histoire, étendu à tous les types d'expression et de

⁷ Voir *infra*, p. 28.

création, en intégrant dans son champ d'application les nouvelles techniques, au fur et à mesure de leur apparition. Depuis 1537, il s'est progressivement étendu aux estampes, aux photographies, aux vidéogrammes, ou encore à l'édition électronique, en 1992. La masse de texte, d'images, de sons, de vidéos, véhiculée par le Web méritait bien d'être prise en compte.

Pourtant, dès que l'on parle du dépôt légal du Web, un certain nombre d'interrogations ne manquent pas de surgir :

- En premier lieu, une des principales exigences du dépôt légal « traditionnel » ne peut pas être honorée : l'exhaustivité. La quantité d'informations, la difficulté de récupérer le Web profond l'interdisent. Il s'agit donc de faire une sélection. Mais selon quels critères ? La question est de savoir si on doit penser cette sélection en terme de politique documentaire, c'est-à-dire acquérir l'intéressant et le rare, ou si on décide de suivre des principes similaires à ceux de l'archivistique, c'est-à-dire faire un échantillonnage le plus large et le plus représentatif possible de l'ensemble du Web (en se refusant donc à distinguer le « bon » du « mauvais »).
- Hésitation entre deux approches. D'une part, la constitution de collection, ce qui implique un travail de veille, et un catalogage de qualité. D'autre part, une approche automatisée, qui privilégie la quantité sur la qualité, et qui reporte à plus tard la question de l'indexation.
- Enfin, comment définir le domaine national ? Selon la loi, tout ce qui est diffusé en France est concerné par le dépôt légal ; or l'ensemble du Web mondial est accessible depuis la France. Une thèse moins maximaliste – et techniquement plus raisonnable – préfère ne considérer que les sites hébergés en France, notamment les 400 000 noms de domaines « .fr » déposés à l'AFNIC⁸. Enfin, que dire des sites francophones, dont l'intérêt pour les internautes et les chercheurs français ne fait aucun doute ?

2.2.2. Les premières expérimentations

Des tentatives de capture massive de sites Web traitant d'un événement ont d'ores et déjà été menées. Pour les élections présidentielles de 2002, du 15 février au 15

⁸ Voir définition en annexe.

juin 2002, 1900 sites différents ont été saisis en 6300 captures. L'archive constituée représente un volume de 530 Gigaoctets et 12 millions de fichiers. L'expérience a été renouvelée pour les élections régionales et européennes de 2004, en affinant les critères et en ciblant mieux l'effort : les captures ont été de meilleure qualité et les volumes saisis moins importants.

2.2.3. Le projet de loi relatif au dépôt légal

Au titre IV du projet de loi sur les droits d'auteur et droits voisins dans la société de l'information, le législateur prévoit d'étendre le champ du dépôt légal aux signes, signaux, écrits, images, sons ou messages de toute nature qui font l'objet d'une communication au public par voie électronique. Cette mesure est présentée comme une exception à la législation sur le droit d'auteur⁹, qui permettra à la Bibliothèque nationale de collecter des œuvres sur le Web, de les conserver et de les communiquer à des lecteurs dûment accrédités, sans demander au préalable une quelconque autorisation.

2.2.4. Collecter

Trois modes d'acquisition ont d'ores et déjà été envisagés à la Bibliothèque nationale de France.

2.2.4.1. Sélection de sites.

Le DBN a établi un réseau de correspondants DL Web dans tous les départements de la BnF, qui doit expérimenter l'aspect « qualitatif » de la collecte, en proposant un certain nombre de sites que la BnF devrait collecter en priorité. Ce réseau est constitué de volontaires, sensibilisés aux enjeux de la recherche d'information sur Internet. C'est pourquoi on y a, le plus souvent, retrouvé les personnes chargées d'alimenter les signets. Cette initiative fut un succès : alors qu'on s'attendait à un millier de sites, une liste de 4 200 sites a été établie. Néanmoins, certaines difficultés ont été constatées. La notion de « sites représentatifs de leurs domaines », que les correspondants devaient recenser, a été différemment appréhendée. Certains départements ont proposé des sites très diversifiés

⁹ Voir *infra*, p. 32.

(regroupant aussi bien des pages personnelles que des sites institutionnels), tandis que d'autres ont travaillé dans une véritable politique d'acquisition, privilégiant des sites de qualité, dont les informations sont bien vérifiées. Le nombre de sites proposés était par ailleurs parfois disproportionné entre les départements.

Les sites ainsi sélectionnés vont faire l'objet d'un « *focus crawl* », c'est-à-dire d'une collecte ciblée. A l'issue de ce travail, les groupes de correspondants seront réunis à nouveau pour évaluer la pertinence des résultats. Ainsi, ce travail répondait à plusieurs objectifs. D'une part, expérimenter et évaluer une méthode de collecte ciblée. D'autre part, sensibiliser les agents de tous les départements de la BnF aux enjeux du dépôt légal du Web.

2.2.4.2. Dépôt

De surcroît, le dépôt de sites Web devrait être spécifiquement prévu par la loi. D'une part, si la BnF veut récupérer un site, alors que son propriétaire en interdit l'accès aux moteurs de recherche du DL Web, la BnF pourra invoquer le « dépôt requis », pour contraindre le propriétaire à lever tout obstacle. D'autre part, des producteurs qui souhaiteraient verser leur site à la BnF, s'apercevant qu'ils n'ont pas été récupérés par le DL Web, pourront solliciter la BnF pour un dépôt spontané. Depuis avril 2005, le DNB collabore avec le Journal Officiel pour mettre en place un dépôt volontaire : tous les soirs, à 22h50, le JO du jour est collecté par un robot de la BnF. Le JO s'assure ainsi la pérennité de ses archives. Pour la BnF, c'est l'occasion d'expérimenter l'automatisation de la collecte et du contrôle qualité, sur un corpus relativement restreint.

2.2.4.3. Capture automatisée par robot

Les « *crawls* larges », c'est-à-dire des collectes couvrant l'ensemble du Web français, peuvent être assurés par des moteurs de recherche capables de parcourir Internet en passant de sites en sites par l'intermédiaire des liens hypertexte. Le contenu de chaque site visité est capturé sur un serveur, et tous les liens hypertexte sont exploités. Néanmoins, la BnF n'a pas les moyens techniques nécessaires pour effectuer elle-même ces *crawls* larges. La solution trouvée a été d'établir un partenariat avec Internet Archives, fondation créée dans le but d'archiver l'ensemble du Web, et qui a commencé son activité dès 1996.

L'outil de recherche utilisé a été développé par le consortium IIPC (International Internet Preservation Consortium), qui regroupe 12 bibliothèques nationales, ainsi que Internet Archives, dans le but de préserver les contenus d'Internet. Il s'agit du robot Héritrix. Ce dernier, malgré sa puissance, présente encore des défauts : ainsi, il ne sait pas gérer les *frames*¹⁰. C'est pourquoi la BnF s'est associée à la British Library pour produire un « *smartcrawler* », qui aurait encore plus de paramètres, réduisant par là le bruit et le silence.

La dernière commande de *crawl* de la BnF à Internet Archives a été hybride. Elle associe un *crawl* large (l'ensemble du domaine national) à un *crawl* ciblé sur les 4 200 sites sélectionnés par la BnF (soit 40 000 URL¹¹). Le *crawl* va durer 6 semaines en tout, la liste ayant été envoyée le vendredi 7 octobre. Après le *crawl*, un temps considérable sera employé, sur place, au contrôle qualité et à l'indexation. C'est en janvier que devrait être envoyé à Paris le serveur contenant les données. En février, les résultats du *crawl* ciblé pourront être analysés avec les correspondants DL Web.

2.2.5. Traiter

2.2.5.1. Contrôle qualité

A l'arrivée du serveur, un contrôle qualité est effectué par le DBN. Un tel travail peut sembler, sur un total estimé de 2,2 milliards d'URL, impossible. Toutefois, des opérations simples, automatisées, peuvent être accomplies : contrôle de réception (vérification du nombre de données reçues par rapport aux données demandées,...), mais aussi des vérifications plus « documentaires » sur la qualité (regarder jusqu'à quelle profondeur le robot a pu collecter, quelles redirections il a utilisées...).

2.2.5.2. La conservation

Les ressources du DL Web seront intégrées dans un système d'archivage global¹². Ce magasin sera au format ARC, spécifiquement destiné à l'archivage de données.

¹⁰ Voir définition en annexe.

¹¹ Voir définition en annexe.

¹² Voir *infra*, p. 28.

Il a été notamment décidé que l'unité documentaire serait le site Web (on aurait pu en juger autrement : choisir la page ou encore le fichier, comme unité documentaire). En effet, on a été jugé que l'essence même du Web, c'était la structure intertextuelle : il n'est possible d'évaluer un contenu qu'en voyant la façon dont il se place au sein des autres contenus.

2.2.6. L'accès

Une réflexion a été menée sur les outils d'accès. Le seul qui est disponible aujourd'hui est la « Wayback machine », que l'on peut utiliser notamment sur le site d'Internet Archives¹³. Son défaut est qu'il suppose qu'on dispose par avance de l'URL que l'on veut consulter. Par conséquent, il faut penser à des outils de recherche et de consultation plus intelligents. C'est pourquoi le consortium IIPC a développé l'outil Wera, qui fait de la recherche en texte intégral. D'autres expérimentations sont en cours : recherche en langage naturel, classement en fonction de la notoriété...

Toutefois, on ne peut pas envisager d'accès tant que la loi sur le dépôt légal du Web n'a pas été votée. Après cette date, les lecteurs accrédités en espace recherche auront accès à un corpus regroupant les résultats des *crawls* par Internet Archives, les sites traitant des élections, et le Journal Officiel.

Ainsi, le modèle conceptuel du fonctionnement du DL Web est validé. Il va falloir passer à une seconde étape, dans la perspective du passage de la loi : la stabilisation du modèle et sa mise en production, c'est-à-dire l'établissement d'une véritable organisation, stable et étoffée, autour du DL du Web.

3. Travailler au DBN

3.1. Moyens humains et financiers

¹³ INTERNET ARCHIVES. *Internet Archives : Wayback Machine*. [Consulté le 9 décembre 2005]. Disponible sur : <<http://www.archive.org/>>

Depuis 2004, les effectifs du DBN ont été considérablement renouvelés et étoffés. Si un agent, responsable du DL du Web, est parti, quatre personnes, deux pour le pôle DL Web, deux pour le pôle bibliothéconomie numérique, ont été recrutés. Ainsi, le département regroupe, outre le chef du département, sept agents : quatre dans le pôle document numérique, trois pour le pôle DL Web. Néanmoins, un informaticien, rattaché au DSI, travaille à plein temps sur les projets du DL Web.

Fonction	Catégorie
Directrice du département	Contractuelle, cat. A
Pôle bibliothéconomie numérique	
Questions de normalisation	Contractuel, cat. A
Accès et services	Conservateur
OAI et métadonnées	Conservateur
Publications en ligne	Conservateur
Pôle DL Web	
Pilotage du projet	Conservateur
Spécifications fonctionnelles	Contractuelle, cat. A
Contrôle qualité	Contractuelle, cat. A
Développements informatiques	Informaticien, contractuel

La proportion importante de contractuels dans le service s'explique par la nouveauté des fonctions occupées par les agents du DBN. La tendance actuelle est toutefois au recrutement de professionnels du monde des bibliothèques, ce qui favorise le travail avec les autres départements dans le cadre de projets transverses. D'autre part, la mise en production du DL du Web devrait entraîner, à terme, le recrutement de personnel de catégorie B ou C.

Si l'on excepte les dépenses liées à la rémunération des agents, le budget du département monte à 250 000 € par an. Il est principalement absorbé par les dépenses liées au dépôt légal du Web : facturation du *crawl* par Internet Archives, dépenses de développement de logiciels dans le cadre de la participation au consortium IIPC.

3.2. Le travail au quotidien

Une grande part du travail est consacré à la veille : visite des sites Internet d'institutions publiques, lecture de rapports, documentation technique (aussi bien sur les nouveaux formats de métadonnées, les nouvelles techniques de publication Internet, ou encore les classifications...). Pour un des agents, ce travail de veille alimente un blog sur Internet et la bibliothéconomie. Ce blog assure trois fonctions principales : formaliser le travail de veille, diffuser les informations en interne, particulièrement aux autres membres du DBN, et en externe¹⁴.

Les agents travaillent la plupart du temps par « projets », décidés de leur propre mouvement, ou à la suite de sollicitations venant d'autres services. Ce mode de travail incite à fonctionner en binôme. En effet, il est souvent nécessaire d'associer les compétences, parfois très pointues, de chacun. La diversité des projets comporte un risque : un éclatement des objectifs du DBN, qui entraînerait une perte de la vision d'ensemble, et une dispersion de la force de travail. C'est pourquoi de fréquentes réunions de service sont nécessaires pour exposer à chacun les orientations actuelles du DBN, et éviter que des réflexions se fassent de façon parallèle, sans concertation entre les agents. La taille réduite du service permet une bonne circulation des informations.

Le DBN est en permanence appelé à travailler en collaboration avec d'autres départements, particulièrement le pôle Bibliothéconomie numérique, qui n'est pas producteur ou gestionnaire, ni de collections, ni d'applications logicielles propres. Des collaborations permanentes sont établies avec le DSI, l'ABN (pour la normalisation et les métadonnées), les départements de la conservation et de la coopération (pour Gallica). Des collaborations ponctuelles peuvent être établies avec des départements de la DCO (expérimentation des *crawls* ciblés, gestion des droits...), le département de la reproduction, la délégation à la diffusion culturelle (pour l'intégration de leurs ressources numériques grâce au protocole OAI...).

En outre, les agents du DBN sont souvent sollicités pour des formations (formation à l'OAI, à la préservation des données numériques) ou à des préparations de concours (préparation des chartistes au concours de l'ENSSIB, préparation Médiadix au concours de conservateur).

¹⁴ Il faut néanmoins souligner que ce blog, s'il est régulièrement consulté par des agents du DBN, est une initiative personnelle et informelle, qui ne relève pas du travail effectué à la BnF. Les propos qui y sont tenus n'expriment en aucun cas la position officielle de la BnF, ni du DBN.

Enfin, certains agents participent à des plages de service public. C'est l'occasion, pour les membres d'un département essentiellement tourné vers le « virtuel », de rencontrer le public de la bibliothèque.

Ainsi, le DBN peut sembler un département atypique, aussi bien en raison de sa taille que du mode de travail de ses agents, qui fonctionnent par projets, et n'assurent pas de production courante. Néanmoins, les collaborations établies avec d'autres départements assurent son intégration au sein de l'ensemble de la BnF. Ainsi, une coopération suivie a été mise en place avec tous les départements impliqués dans la production de documents numériques.

Partie 2 : la chaîne de traitement des documents numériques à la BnF

L'un de mes objectifs, lors de ce stage, était de découvrir la chaîne de traitement du document numérique, en prenant principalement l'exemple de Gallica, la bibliothèque numérique de la BnF. C'est pourquoi j'ai voulu rencontrer l'ensemble de ses acteurs, et découvrir ainsi un mode de travail plus orienté vers la production. Les observations qui sont présentées ici sont essentiellement issues des discussions que j'ai eues avec des acquéreurs de Gallica (département de la coopération), des chargées du catalogage et du contrôle qualité de Gallica (département de la conservation), de l'adjointe du chef de service numérisation (département de la coopération). Faute de disponibilité de leur part, il ne m'a pas été possible de rencontrer des représentants du DSI. C'est pourquoi les aspects concernant le DSI m'ont été expliqués par ma responsable de stage.

J'ai enfin eu l'occasion d'assister à une réunion de membres du DBN avec un responsable du département de la reproduction. Ce dernier nous a expliqué la procédure de sélection et de catalogage des documents présents dans la banque d'images. Ces connaissances étaient nécessaires pour permettre au DBN d'établir les bonnes règles de *mapping*.

Cette rencontre m'a permis de prendre concrètement conscience que Gallica n'était pas la seule productrice de ressources numériques à la BnF. La diversité de ces ressources n'est pas un phénomène négligeable. C'est ce qui rend plus intéressant encore la volonté de toutes les intégrer dans un seul système d'archivage.

1. Gallica

Gallica est née en 1997. Elle peut aujourd'hui se prévaloir de plus de 77 000 imprimés en mode image, dont 17 000 fascicules de périodiques, 1 500 documents en mode texte, 80 000 images, et plus de 30 heures de sons. Cette bibliothèque numérique n'a pas d'équivalent sur le Web.

1.1. Sélection et acquisition

Quatre personnes, au sein du service « pôles associés-Gallica », qui dépend du département de la coopération, sont chargées de la sélection des documents dans leur domaine de connaissance. Elles s'occupent aussi de la recherche et de l'acquisition des ouvrages qu'elles ont sélectionnés.

1.1.1. Critères de sélection.

- Intérêt intellectuel. Ce critère est obtenu grâce à une intense documentation sur le thème concerné. Des ouvrages de référence (dictionnaires thématiques et biographiques, bibliographies, ouvrages de fond et de référence sur le sujet) sont consultés pour appréhender l'importance et la représentativité d'un auteur et d'une œuvre.
- Respect de la charte documentaire, qui fixe les grands principes de Gallica. Les livres qui y sont contenus doivent être difficilement accessibles, fragiles, et importants pour la recherche. Ainsi, nous sommes dans une optique de numérisation de diffusion, pas de sauvegarde.
- Prix du livre (pour les achats auprès de libraires spécialisés).
- Date d'édition (sélection de documents passés dans le domaine public¹⁵).
- Aspects matériels : le livre doit être en bon état, avec de larges marges.
- Suggestions de lecteurs individuels sur la « base de suggestion ».
- Partenariats : il s'agit le plus souvent de projets ponctuels, voire de demandes de chercheurs individuels. Par exemple, un projet de numérisation de rapports de congrès ouvriers a été mené avec le CODHOS (Collectif de documentation sur l'histoire ouvrière et sociale). Dans des cas semblables, les tâches sont ainsi réparties : le partenaire fait un répertoire de toutes ses ressources, que Gallica se charge de numériser. Ainsi, au premier l'expertise scientifique, le choix et la bibliographie, au second la mise en œuvre technique, et la coordination sur les problèmes de droits (c'est par exemple la BnF qui rédige les conventions).

¹⁵ Voir *infra*, p. 32.

1.1.2. Origine des ouvrages

- Microformes : c'est la solution la plus avantageuse. En effet, aucun travail sur l'ouvrage original n'est demandé, et la numérisation est rapide.
- Livres obtenus grâce à un partenariat.
- Livres issus de la collection de l'ancien centre de prêt de Versailles, fermé dans les années 1990. Ce centre dépendait de la BnF, et désormais ses ouvrages font doublon avec ceux du dépôt légal.
- Antiquariat : c'est l'origine la plus répandue. En effet, pour être numérisés par les prestataires, les livres sont massicotés (le coût d'une numérisation page à page étant le plus souvent supérieur à celui de l'achat du livre lui-même). Dans ce cas, à l'issue du contrôle qualité, les ouvrages sont envoyés au pilon.
- Demande d'ouvrages dépendant de la DCO. Cette procédure est assez rare. En effet, il n'est pas question de massicoter ces ouvrages. Cette procédure est employée lorsqu'un document a été rejeté par le prestataire, et que sa présence au sein de Gallica est indispensable (il fait partie d'une collection, par exemple).

Ceci explique la prédominance du XIX^e siècle dans Gallica : les livres du XX^e sont couverts par les droits d'auteur, tandis que ceux du XVI^e-XVIII^e sont trop chers à l'achat. Toutefois, l'usage de microformes permet de nuancer cette tendance.

1.1.3. Budget

Chaque coordinateur Gallica dispose d'un budget de 10/11 000 € par an pour ses achats. La dépense moyenne par achat est de 100 €. L'achat est intéressant à partir d'un *ratio* de 30 centimes d'euro par page environ. A partir d'une limite maximum de 50 centimes, il est préférable de faire de la numérisation page à page. Les sélections d'achat se font à partir des catalogues de libraires, qui sont envoyés directement à la BnF, ou en ligne (un portail comme Galaxidion donne accès à divers sites de libraires d'ancien).

Après avoir été sélectionnés, les ouvrages sont confiés au département de la conservation, qui prépare et contrôle la numérisation effectuée par des prestataires extérieurs.

1.2. Numérisation

1.2.1. Les filières de numérisation

Il en existe trois, correspondant chacune à un marché avec un prestataire différent :

- Papier : massicoté (80 000 pages / mois) et relié (16 000 pages / mois).
- Microfilm. A la fois microfiches (20 000 pages / mois, surtout pour de petits volumes) et microfilms (26 000 pages / mois, plutôt pour de grands volumes).
- Lots d'images.

Enfin, les tables des matières sont saisies par un autre prestataire.

1.2.2. Vérification et catalogage

Il s'agit de vérifier, livre en main (alors que les acquéreurs travaillent sur des catalogues), le nombre de contributeurs : illustrateurs, préfaciers..., afin d'être sûr qu'aucun droit d'auteur ne peut s'appliquer. Ensuite, on catalogue à partir de la notice de BN-Opale Plus. Après cette étape, l'ouvrage est expédié au prestataire.

1.2.3. Contrôle qualité.

Les documents numérisés, renvoyés par le prestataire, sont vérifiés après réception. L'agent chargé du contrôle dispose d'un logiciel de visualisation qui lui indique automatiquement quelles pages contrôler : les premières (titre...), les dernières, les tables de matières, et d'autres, aléatoirement choisies par le système de contrôle. Au début du marché, 25% des pages doivent être contrôlées ; après, on passe à 10%. L'agent doit signaler tout problème rencontré, et il est souvent nécessaire de vérifier sur l'ouvrage original pour identifier le responsable. Les problèmes les plus fréquents sont d'ordre technique (ondulation des pages, mauvaise luminosité...) ou liés à une mauvaise pagination.

Certains défauts, peu importants, sont laissés en l'état. Toutefois, on insère alors un logo indiquant la nature du défaut, pour en informer le lecteur.

Lorsque les documents ont été validés par le contrôle qualité, des lots de migration sont faits sur les serveurs du DSI, au début de chaque mois.

1.3. Migration et consultation

1.3.1. La migration sur les serveurs du DSI

Deux types de données sont disponibles dans les CD renvoyés par le prestataire : les images, et un fichier XML nommé « refnum ». Ce fichier très important indique, pour les imprimés, la correspondance entre le numéro de l'image de texte, et le numéro de page de l'ouvrage initial. Ceci sera nécessaire pour que le lecteur se repère à l'avenir. Pour les lots d'images, il contient les légendes des images.

Les images de texte sont au format TIFF noir et blanc, qui offre l'avantage d'une compression très forte, mais également totalement réversible. Ainsi, cela permet à la fois pérennité des informations, et légèreté des données. Les images en revanche sont en JPEG, tandis que la presse est en TIFF niveaux de gris.

Lors du passage sur les serveurs du DSI, les données fournies par le prestataire vont passer par une « moulinette », qui va les orienter sur deux réservoirs :

- Le réservoir des images numérisées.
- Un réservoir contenant des métadonnées : on y trouve le « refnum », des informations techniques (poids, taille des fichiers), des informations sur le déroulement de la migration...

Au même moment, des changements s'opèrent dans le catalogue BN-Opale Plus : la notice de l'ouvrage concerné se dote d'une vignette « document numérisé ». D'autre part, la notice BN-Opale Plus passe dans une base MySQL qui permet d'effectuer des recherches dans Gallica.

1.3.2. L'interface de consultation

Lorsqu'un usager demande une ressource numérique à partir de Gallica, deux logiciels, communiquant entre eux par une série de requêtes, vont être sollicités.

Le premier est le « médiateur ». C'est lui qui va puiser dans les deux réservoirs (celui contenant les images, et l'autre les métadonnées) et dans BN-Opale Plus, pour mettre en relation images, métadonnées, et notices. Il transmet les données au « visualiseur », qui permet leur consultation sur l'écran de l'utilisateur¹⁶.

2. Les bases du département de la reproduction

Deux bases existent : la base Daguerre (base de production), et la banque d'images (base de consultation). A l'heure actuelle, il y a 85 000 images visibles dans la banque d'images, sur les 135 000 que compte Daguerre. En effet, il n'est pas possible d'afficher des images dont on craint qu'elles soient encore sous droit. Néanmoins, une partie des 70 000 images non accessibles le sont par prudence, par défaut de renseignements. Le rythme de production est de 1 000 images par mois.

2.1. Alimentation de la base.

Les documents sont toujours photographiés à la demande d'un client, qui paie la prise de vue. Si le format est assez grand, le cliché est intégré à la banque d'images, qu'il soit numérique ou argentique (dans ce cas, il est numérisé). Néanmoins, la priorité du département est de réduire les délais de traitement des demandes. C'est pourquoi, faute de temps, la numérisation rétrospective de clichés argentiques n'est pas toujours effectuée.

2.2. Description et mise en ligne

Ensuite, les images sont transférées sur un des deux serveurs du département de la Reproduction (à Richelieu ou Tolbiac). Les légendeurs du département vont alors les décrire, les ayant sur leur écran, en qualité maximale.

Ensuite, la notice est envoyée au département des collections concerné – essentiellement les départements spécialisés –, qui valide les descripteurs choisis. Après validation de ce département, l'image est mise en ligne.

¹⁶ Voir schéma en annexe.

La procédure d'obtention de ressources pour un usager est la même que celle de Gallica. Lorsqu'une ressource, identifiée par une recherche dans la banque d'images, est demandée par l'utilisateur, c'est le même visualiseur qui est sollicité. Toutefois, ce visualiseur présente à l'utilisateur une interface différente, pour qu'on ne puisse pas confondre Gallica et la banque d'images.

3. L'archivage pérenne et le magasin numérique

A terme, toutes les données numériques produites par la BnF (c'est-à-dire originellement produites sous forme numérique ou numérisées par la BnF) devront être regroupées au sein d'un vaste ensemble, dont le nom n'est pas encore choisi : système d'archivage, entrepôt ou magasin numérique. Sa mise en place est l'objectif du projet SPAR (Système de Préservation des ARchives). Un groupe de travail sur l'archivage des données numériques s'est formé dans cet objectif. Les enjeux sont énormes : il s'agit de stocker, dans un même ensemble, des données sous des supports variés ; de garantir leur pérennité, malgré les obsolescences rapides des formats ; et de contrôler à l'avenir l'accès à ces données. C'est pourquoi « la conservation à long terme des données numérique est un des défis les plus redoutables auxquels sont aujourd'hui confrontées les institutions chargées de constituer notre mémoire collective »¹⁷.

3.1. La variété des producteurs de données numériques

- Département de l'Audiovisuel : il présente un cas à part, en assurant une grosse production documentaire. Tous les supports analogiques sont en train de passer au numérique, ce qui représente à terme des masses gigantesques : pas moins de 460 Téraoctets d'ici 10 ans. Ce passage au numérique est rendu nécessaire par la disparition des lecteurs pour les supports analogiques (magnétoscopes,

¹⁷ **Jacqueline MELET-SANSON, Daniel RENOULT (dir.).** *La Bibliothèque nationale de France : collections, services, publics.* Paris : Électre – Cercle de la Librairie, 2001, p. 209.

lecteurs de disques vinyles...). Un programme de conversion sur 5 ans, commencé en 2002, doit assurer ce vaste changement de support.

- Département de la conservation : Gallica. Pour l'instant, les volumes ne sont pas importants, mais ils vont considérablement augmenter, notamment à cause de la numérisation de la presse. De même, ce département va assurer la numérisation de préservation¹⁸.
- Dépôt légal : des éditeurs, notamment de presse, peuvent préférer déposer une version numérique plutôt qu'une version papier.
- Dépôt légal du Web. Dans dix ans, les estimations prévoient un volume total de 800 Téraoctets, soit près de la moitié des besoins de stockage du système d'archivage.
- Département de la reproduction (banque d'images), délégation à la diffusion culturelle (Expositions virtuelles, Dossiers pédagogiques), délégation à la communication (photothèque)...
- Service des archives de la BnF. Un vaste programme, piloté par la mission pour la gestion de la production documentaire et des archives, prévoit que tout document numérique produit par un agent de la BnF doit entrer, pour être stocké sur les serveurs, dans une typologie (notes de service, guides d'utilisation...). En fonction de son type, chaque document se voit attribuer une durée de vie. Le courrier reçu par voie postale sera lui-même numérisé. Au bout d'un certain temps, en fonction de la réglementation sur les archives en vigueur, le public aura accès à ces documents.

3.2. Difficultés rencontrées

3.2.1. Problèmes de supports et de formats

Un important travail d'uniformisation est à mettre en œuvre. Jusqu'à maintenant, chaque département assurait la pérennité avec ses propres méthodes, usant de supports différents (disques de verre, bandes magnétiques) et de divers formats : TIFF, JPEG, MPEG, HTML... Les choix étaient effectués en fonction des moyens et des compétences de chaque service.

¹⁸ Voir *infra*, p. 34-35.

Les formats sont appelés à vieillir et disparaître. C'est pourquoi il va falloir effectuer une veille attentive pour surveiller les modifications de tous ceux qui pourront être utilisés. En outre, on utilisera des logiciels à même de détecter automatiquement quels supports, vieillissant, méritent d'être remplacés, et d'effectuer automatiquement la migration de l'ancien support au nouveau.

3.2.2. Métadonnées

Pour garantir l'accès aux documents, ces derniers devront être correctement indexés, c'est-à-dire bénéficier de métadonnées fiables. Par exemple, il est nécessaire que chaque ressource dispose d'un identifiant unique et pérenne. C'est pourquoi tous les documents entrant dans le système feront l'objet d'un contrôle qualité – sommaire et automatisé –, même ceux provenant de l'extérieur.

3.3. Modèle cible

L'objectif final est d'obtenir un système d'archivage conforme à la norme OAIS¹⁹. Ce modèle, devenu une norme ISO (ISO 14721 : 2002), définit un vocabulaire et un ensemble de concepts indispensables à la compréhension de la problématique de la pérennisation des données numériques. Il définit et se préoccupe de cinq grandes fonctions : production, versement, stockage, préservation et utilisation²⁰. Ainsi, il est non seulement nécessaire de réfléchir aux ressources numériques produites et aux services producteurs, mais aussi aux besoins futurs des utilisateurs.

Actuellement, la norme OAIS est largement tributaire des réflexions des scientifiques, notamment dans le cadre de la recherche spatiale. Il s'agissait de traiter des images satellites de grande qualité, très volumineuses.

En France, c'est le CNES (Centre national d'études spatiales, à Toulouse), qui a mis au point la première architecture d'archivage selon le modèle OAIS. Actifs depuis 10 ans, ils archivent désormais pour le compte d'autres institutions, comme

¹⁹ Voir définition et schéma en annexe.

²⁰ L'OAIS est très bien expliqué dans Sylvie DALBIN. « Compte rendu de la journée d'études : La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ? ». *Documentaliste - Sciences de l'information*. [en ligne]. 2003, vol. 40, n° 3, p. 226-231. [Consulté le 23 novembre 2005]. Disponible sur <http://www.adbs.fr/uploads/journees/2293_fr.php>.

le Bureau de recherches géologiques et minières. Ils ont atteint un total de 400 Téraoctets stockés en mai 2005.

Moins avancée, la BnF en est à une phase d'évaluation, et d'analyse des risques. Des appels d'offre sont en préparation, d'autres – notamment pour le stockage – ont déjà été lancés. D'autres institutions publiques sont intéressées par ce modèle, comme le Centre informatique national de l'enseignement supérieur (CINES) ou les Archives Nationales, qui suivent avec attention les résultats de la BnF.

Ainsi, la question du document numérique, à la BnF comme dans les autres institutions publiques, dépasse largement celle de la numérisation. C'est la conservation à long terme du patrimoine, mais aussi, plus prosaïquement, de la mémoire administrative d'un établissement, qui est en jeu. Aucune réflexion sur le stockage ne peut en outre se passer d'une analyse sur les divers usages et accès à ces documents, dès aujourd'hui et pour le futur. Or, ces accès sont l'objet d'une réglementation précise, en permanente évolution. Pour les archives, c'est la loi de 1979, récemment codifiée dans le code du patrimoine, qui s'applique. Pour les « œuvres de l'esprit » en revanche, une réflexion s'impose, qui doit particulièrement prendre en compte la législation sur la propriété intellectuelle.

Partie 3 : activités effectuées par le stagiaire

1. La réflexion sur le statut juridique du document numérique

1.1. La législation confrontée au défi du numérique et de l'Internet²¹

1.1.1. La législation en vigueur

En France, le droit d'auteur s'est toujours développé comme le compromis entre trois exigences : la revendication de la part des auteurs à une juste rémunération, le respect de leur œuvre, et l'intérêt du public pour la diffusion la plus large possible.

Le droit d'auteur s'est ainsi subdivisé en deux catégories :

- le droit moral, attaché à la personne, perpétuel, inaliénable et imprescriptible. Il est composé de quatre prérogatives essentielles : le droit à la paternité (droit au respect du nom ou droit pour l'auteur d'être cité), le droit de divulgation (décider du moment où l'œuvre est divulguée au public et sur quel support), le droit au respect de l'œuvre (droit de défendre l'œuvre contre toutes les atteintes qui pourraient y être portées), le droit au repentir ou droit de retrait de l'œuvre.
- le droit patrimonial. En vertu de ce droit, l'auteur est titulaire d'un droit de reproduction (c'est-à-dire la fixation matérielle de l'œuvre par tous procédés qui permettent de la communiquer au public de manière indirecte) et de représentation (communication de l'œuvre au public par un procédé quelconque).

Contrairement au droit moral, le droit patrimonial est cessible, dans le cadre d'utilisations commerciales. Il n'a qu'une durée limitée, fixée par la loi. Lorsqu'il cesse, les œuvres tombent dans le domaine public.

²¹ Une présentation détaillée de l'ensemble des droits d'auteur, et du contexte technique, économique et juridique qui les fait évoluer, n'est pas l'objet, faute de place, d'un tel rapport de stage. Une présentation succincte, complète et très à jour de cette question peut être l'article de **Michèle BATTISTI**. « Le droit d'auteur, un obstacle à la liberté d'information ? », *Bulletin des bibliothèques de France* [en ligne]. 2004, n°6, p. 31-35. [Consulté le 23 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

Outre le droit d'auteur, d'autres types de droit peuvent être attachés à des documents numériques. En premier lieu, il s'agit des « droits voisins », reconnus à certains auxiliaires de la création littéraire et artistique : artistes-interprètes, producteurs de phonogrammes ou de vidéogrammes, entreprises de communication audiovisuelle.

D'autre part, le droit à l'image, le droit au respect de la vie privée, s'appliquent aussi bien pour les documents numériques que pour leurs équivalents analogiques.

1.1.2. Le défi présenté par les nouvelles technologies.

Les nouvelles technologies ont démultiplié la capacité de produire et de diffuser des documents numériques. Les pages personnelles, les dépôts d'archives ouvertes, les versions en ligne des principaux quotidiens français d'information en sont autant d'exemples. D'autre part, le remplacement de l'analogique par le numérique dans les domaines de l'image, du son ou de la vidéo, permettent de reproduire les contenus pratiquement sans altération.

Dès lors, les exemples d'infraction aux règles du droit d'auteur sur Internet sont légion, puisqu'il est aisé de reprendre à son propre compte des textes qui y sont diffusés, ou encore de réutiliser sur son site personnel des clichés sans avoir l'autorisation de la personne représentée. L'exemple le plus connu, car économiquement le plus préjudiciable, est celui des réseaux « *peer to peer* », c'est-à-dire d'une technologie permettant l'échange de documents entre internautes, et qui est abondamment utilisée pour télécharger, de façon gratuite et illicite, des musiques ou des films. Des utilisateurs ont été condamnés. Une autre initiative a eu un grand retentissement : celle de Google Print, qui a décidé de mettre en ligne des extraits numérisés d'ouvrages sous droit, sans l'accord de leurs ayants droit. Deux procès viennent d'être intentés contre Google, l'un par des éditeurs, l'autre par une société d'auteurs²².

²² Claudine MULARD. « La bibliothèque numérique de Google dénoncée par cinq grands éditeurs américains ». *Le Monde*, 25 octobre 2005.

1.1.3. La réaction : renforcement de l'arsenal juridique et technique.

Un nouvel équilibre entre diffusion et protection devait en conséquence être trouvé. C'est le second aspect qui tend désormais à être privilégié, en raison notamment du lobbying mené par les producteurs (éditeurs, producteurs de musique ou de films). Cette tendance s'est exprimée aussi bien dans les traités de l'Organisation mondiale de la propriété intellectuelle (OMPI), en 1996, que dans le *Digital Millenium Copyright Act* (DMCA) américain, en 1998, et dans la directive européenne sur le droit d'auteur et les droits voisins dans la société de l'information (DADVSI), en 2001. Cette directive devrait être transposée en droit français par le projet de loi relatif aux droits d'auteur et aux droits voisins dans la société de l'information, qui sera présentée à l'Assemblée Nationale en décembre 2005, et étudiée par le Sénat en janvier 2006²³.

L'un des principaux objectifs de ces textes est de renforcer juridiquement les protections techniques qui pourront être installées par les producteurs. En effet, une nouvelle série d'outils voit le jour, appelés DRM (Digital Rights Management), outils de gestion des droits numériques.

Néanmoins, la directive européenne établit une liste fermée, mais large, d'exceptions possibles au droit d'auteur. Le projet de transposition français n'en retient que deux, l'une en faveur des personnes handicapées, l'autre en faveur des établissements dépositaires du dépôt légal²⁴.

1.2. L'influence de ce nouveau contexte à la BnF

La transposition de la directive européenne va ainsi avoir de nombreuses conséquences pour la BnF. En premier lieu, le projet de loi prévoit de rendre possible le dépôt légal du Web²⁵. D'autre part, il devrait l'autoriser à exécuter des copies numériques d'ouvrages sous droit, dans le cadre de campagne de conservation préventive (afin d'éviter la manipulation des documents par les

²³ Les références de la directive et du projet de loi sont disponibles dans la bibliographie.

²⁴ Elles viennent s'ajouter aux autres exceptions prévues par la législation antérieure : représentation dans le cercle de famille, copie privée, analyse et courte citation, revue de presse, pastiche et satire.

²⁵ Voir *supra*, p. 15.

lecteurs), et à les diffuser, mais uniquement aux personnes accréditées dans l'espace recherche. Le numérique remplacera ainsi le microfilm.

Le besoin d'un système de gestion des droits des documents numérisés va dès lors se faire sentir. Ce système devra être à même, pour chaque document, de calculer automatiquement les droits d'accès, c'est-à-dire de savoir s'il doit fournir la ressource uniquement à une personne en Rez-de-Jardin, en Haut-de-Jardin, ou le mettre à disposition de tous sur le net. D'autre part, il devra gérer des droits d'usage (autoriser ou interdire le téléchargement, la photocopie...).

Pour décider, le système aura besoin que des métadonnées de gestion de droits soient renseignées. Il s'agit d'informations sur le document lui-même (date de publication, date de mort de l'auteur, présence d'une personne sur une photographie...), pas d'interprétations sur le statut juridique.

Il a été décidé que les métadonnées de gestion des droits seraient modélisées sous la forme d'arbres de décision par types de documents. L'arbre de décision revient à poser un certain nombre de questions sur un document, dans un certain ordre, de façon à déterminer sa communicabilité.

1.3. Travail effectué par le stagiaire

1.3.1. Les arbres de décision

La modélisation de ces arbres de décision était le principal objet de mes travaux au sein du DBN. J'ai travaillé sous la direction de Frédéric Martin, en collaboration avec le service juridique de la BnF. Il fallait prendre en compte les cas les plus complexes pour s'assurer que toutes les métadonnées seraient prévues.

Les arbres étaient élaborés lors des réunions avec le service juridique. J'étais alors chargé de les formaliser, et de les tester en prenant divers exemples de documents qui pourraient être numérisés. Ensuite, les arbres étaient validés par le service juridique au cours de la réunion suivante.

La collaboration avec le service juridique était indispensable pour percevoir certaines particularités du droit d'auteur français. Ainsi, les prorogations de guerre. Le législateur a considéré que les deux guerres mondiales ont été des périodes pendant lesquelles les auteurs n'ont pas pu toucher de rémunérations sur leurs

œuvres. C'est pourquoi ils ont prévu que les œuvres parues avant la Première ou la Seconde guerre mondiale auraient un délai de protection plus long. Par conséquent, dans les arbres de décision, trois cas de figure ont été retenus : publication antérieure à 1919 / entre 1919 et 1948 / postérieure à 1948.

Ce travail a permis d'élaborer six arbres de décision, distingués en fonction des supports²⁶. Ils ont été validés le 23 novembre par une réunion du groupe de travail sur les droits, qui comprend des représentants de la DSR (DBN, conservation, DSI), de la DCO, et le service juridique. A l'issue de cette réunion, il a été décidé de continuer le travail entrepris : modélisation d'autres supports (objets en 3 dimensions, archives numériques...), tests à grande échelle...

1.3.2. L'expression des droits des ressources numériques mises en ligne par la BnF

De surcroît, j'ai été chargé de recenser les différentes façons dont les droits associés aux documents numériques étaient exprimés sur le site Internet de la BnF. En effet, chaque producteur de document les exprime de façons différentes, en fonction de ses traditions et de ses besoins. Ainsi, les Expositions virtuelles indiquent les crédits de chaque cliché, sans déclaration générale. Le département de la reproduction, en revanche, propose une page générale où les redevances d'utilisation pour toutes les images sont tarifées. Enfin, Gallica présente une déclaration générale pour les droits des ressources disponibles sur le site, mais utilise des formulations différentes pour certains de ses dossiers (comme le dossier France/Amérique).

D'autre part, il s'agissait de voir comment d'autres institutions culturelles exprimaient ces mêmes droits. Ainsi, j'ai pu comparer – entre autres – les formulations de la base Atlas du musée du Louvre, de la base Enluminures (qui recense les enluminures des bibliothèques publiques de France), ou encore du CNUM (bibliothèque numérique du CNAM).

L'objectif final était, en accord avec le service juridique, de dégager les informations qui méritaient d'être indiquées – et qui pouvaient l'être à bon droit, en fonction des évolutions de la législation. Il s'agissait également de réfléchir à

²⁶ Les arbres de décision, ainsi que leur « mode d'emploi », qui fut proposé au groupe de travail sur les droits, sont disponibles en annexe.

une possible unification des déclarations, puisqu'une recherche fédérée sur toutes les ressources sera bientôt disponible grâce au protocole OAI. Des pistes ont été dégagées, qui devront être soumises à une prochaine réunion de groupe de travail sur les droits, avant d'aboutir à une formulation définitive. En premier lieu, l'unification des déclarations a été abandonnée, en raison d'une trop grande divergence des objectifs des divers départements. Deuxièmement, il a été reconnu nécessaire de mettre l'accent sur la différence entre le contenant (les sites BnF ou Gallica, protégés au titre du droit d'auteur), et le contenu (les ressources, qui peuvent être dans le domaine public). Troisièmement, une formulation moins restrictive des autorisations, particulièrement pour Gallica, a été souhaitée.

1.3.3. Apport de ce travail

Les discussions ont toujours été très fructueuses. Elles m'ont fait percevoir la complexité du statut juridique d'un document numérique, et comprendre que la prudence était de mise. J'ai également pu prendre conscience que le droit n'était pas un bloc monolithique, mais une série de règles, qui semblent parfois contradictoires, en permanente évolution. Ainsi, la législation européenne avait fait naître une protection spécifique aux bases de données, mais cette protection tend à s'amoindrir... sous l'effet de la jurisprudence de la Cour de justice des communautés européennes. La comparaison entre les différentes sources de droits, l'attention aux évolutions probables de la législation, la prise en compte des missions de la BnF, de la vocation spécifique de chaque département, et enfin des attentes des usagers, formèrent la matière de toutes nos réflexions.

2. Autres activités

2.1. Participation aux activités du service

Au cours de ces douze semaines de stage, j'ai été associé aux différentes activités du service. J'ai notamment participé à diverses réunions : réunions de service,

réunions transverses destinées à débattre d'un sujet (fonctionnement de l'entrepôt OAI ou gestion des droits numériques), réunions de travail dans le cadre des travaux sur les arbres de décision²⁷. Les réunions de service étaient d'autant plus intéressantes qu'elles étaient un moment privilégié où les agents des deux pôles pouvaient exposer à leurs collègues l'avancée de leurs travaux, et où les objectifs de ces pôles étaient débattus.

Le contraste était frappant avec une autre réunion, celle du service de la Bibliographie nationale française livres, qui regroupe plus d'une soixantaine de personnes : le nombre de participants ne permettait pas le même échange d'idées, et la réunion devait être animée par le chef de service²⁸.

Enfin, j'ai eu l'occasion d'assister à une conférence intitulée « La protection des documents et la gestion des droits numériques », organisée par le Centre des hautes études internationales d'informatique documentaire. Cette intervention, et le débat – parfois animé – qui s'en est suivi, ont été l'occasion de constater l'importance de cette question, mais aussi de faire un constat rassurant, puisque l'intervenant s'est montré, au cours de sa présentation, très soucieux de la question du passage des documents dans le domaine public.

J'ai également pu effectuer des tâches plus concrètes, comme le contrôle qualité de manuscrits numérisés fournis par un prestataire. Enfin, j'ai participé à des plages de service public, en Rez-de-jardin (salle K : philosophie et religion) et Haut-de-jardin (salle F : histoire de l'art). Le public du Rez-de-jardin m'est apparu plus autonome, puisque ses questions correspondaient le plus souvent à des problèmes de demande de communication exceptionnelle d'ouvrages, tandis que celui du Haut-de-jardin nous sollicitait plutôt pour des recherches bibliographiques.

2.2. Découverte de la BnF

²⁷ La liste de toutes les rencontres et réunions est disponible en annexe.

²⁸ Nous avons été invités à cette réunion avec un autre membre de notre groupe-projet de l'Essib, qui avait travaillé sur l'adaptation des produits bibliographiques de la BnF aux besoins des utilisateurs. En effet, les principales conclusions de ce projet – notamment l'opinion des bibliothèques interrogées sur le contenu des notices de la BnF – faisaient l'objet d'une présentation lors de la réunion de service. A l'issue de l'exposé, nous avons ainsi pu répondre aux diverses questions posées.

Pour tous les stagiaires de la BnF étaient organisées des rencontres qui avaient pour vocation de nous faire découvrir, au-delà de notre département, l'établissement dans sa globalité. Cinq rencontres ont été organisées :

- budget/ politique documentaire/ acquisitions/ documentation électronique
- traitement des collections (notamment du dépôt légal)
- numérisation / coopération
- services au public sur place et à distance
- Politique d'accréditation / étude des publics.

D'autre part, des visites étaient prévues, comme la visite de la Gestion centralisée des acquisitions, ou celle du département de l'Audiovisuel. Enfin, nous avons pu nous intégrer aux présentations qui étaient organisées pour les agents nouvellement entrés à la BnF. C'est ainsi que nous avons pu visiter le site technique de Bussy, où est entreposée la collection de sécurité de la BnF (désormais close), et où sont installés des ateliers de sauvegarde des documents : restauration, désinfection...

Ces rencontres et visites m'ont permis d'avoir une bonne vision d'ensemble de l'organisation de la BnF (les principales directions...), de ses missions (notamment le dépôt légal, mais aussi la coopération), de ses moyens (budgétaires et humains), de ses publics, et aussi des méthodes d'évaluation qui lui permettent d'estimer la réussite de ses objectifs. Ce qui m'a le plus intéressé, dans les interventions que nous avons pu entendre, c'était d'y retrouver les mêmes objectifs et les mêmes problèmes que ceux qui se posaient au DNB. Le dépôt légal en ligne de la presse quotidienne régionale concerne aussi bien le DNB que le département du dépôt légal ou celui de la coopération. Le projet SPAR a été évoqué au cours de plusieurs rencontres. Les personnes chargées de la documentation électronique ou des signets recherchent, elles aussi, à appliquer au numérique les méthodes pratiquées pour les documents analogiques, tout en en dégageant leurs spécificités.

Conclusion

Pour conclure, je souhaiterais souligner à quel point ce stage a été enrichissant pour ma formation de conservateur. Ce fut l'occasion d'acquérir des connaissances d'ordre juridique et technique, dont l'importance dépasse le cadre des seules problématiques de la numérisation. J'ai pu prendre conscience aussi bien des problèmes juridiques qui se posent aux bibliothèques que de l'importance d'un choix pertinent des métadonnées ou de la pérennité des supports et des formats.

Ce stage m'a ensuite permis de percevoir la diversité des fonctions d'un conservateur, même si c'était sous l'angle particulier des documents numériques : politique d'acquisition (pour Gallica ou pour l'expérimentation du DL du Web), traitement des documents (définition des métadonnées, contrôle qualité), conservation, « accueil » du public (réflexion sur les interfaces), travail avec des prestataires privés, souci systématique d'évaluation.

En revanche, je n'ai guère pu observer une tâche essentielle pour la plupart des conservateurs de bibliothèque : la fonction d'encadrement. En effet, les agents du DBN travaillent de façon très autonome, essentiellement par projets. Ils n'assurent aucune production de documents – cela va bientôt changer pour le pôle du DL du Web. Ce mode de fonctionnement a pour avantage de les inciter à collaborer en permanence avec l'ensemble des autres départements de la BnF. Ceci m'a permis, en sus des rencontres qui étaient organisées, de découvrir le fonctionnement de cet établissement si important au sein du réseau des bibliothèques françaises.

Enfin, ce stage fut l'occasion de percevoir une dimension essentielle du métier de conservateur : la médiation. Les agents du DBN assurent par exemple une interface entre les demandes des bibliothécaires et les contraintes exposées par les informaticiens. Pour cela, ils doivent se situer à un niveau « technico-stratégique », c'est-à-dire être capables d'identifier les objectifs réalisables, de comprendre la technique, sans nécessairement savoir la mettre en œuvre. Le travail effectué avec le service juridique, puis la présentation des arbres de décision aux autres départements, est une autre forme de cette médiation. Ainsi, la qualité principale d'un conservateur est peut-être l'écoute et la compréhension.

Bibliographie

La Bibliothèque nationale de France :

BIBLIOTHEQUE NATIONALE DE FRANCE. *Rapport d'activité 2004.* Paris : BnF, 2005, 76 p.

FRANCE. Commission des affaires culturelles du Sénat. *Rapport d'information à la suite d'une mission d'information chargée d'étudier le fonctionnement de la Bibliothèque nationale de France.* Paris : Sénat, 2000, 93 p.

MELET-SANSON Jacqueline, RENOULT Daniel (dir.). *La Bibliothèque nationale de France : collections, services, publics.* Paris : Electre – Cercle de la Librairie, 2001, 238 p.

STASSE, François. *La véritable histoire de la grande bibliothèque.* Paris : Éd. du Seuil, 2002, 205 p.

Documents numériques :

BIBLIOTHEQUE NATIONALE DE FRANCE. *Charte documentaire de Gallica* [en ligne]. Paris, BnF, 2004. Mise à jour le 13 septembre 2005. [Consulté le 27 novembre 2005]. Disponible sur : http://www.bnf.fr/pages/infopro/cooperation/pdf/Charte_Gallica.pdf.

DALBIN, Sylvie. « Compte rendu de la journée d'études : La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ? ». *Documentaliste - Sciences de l'information* [en ligne]. 2003, vol. 40, n° 3, p. 226-231. [Consulté le 23 novembre 2005]. Disponible sur http://www.adbs.fr/uploads/journees/2293_fr.php.

LUPOVICI, Catherine. « L'archivage de l'Internet à la Bibliothèque nationale de France ». *Matériaux pour l'histoire de notre temps*, 2005, n°79, p. 82-84.

LUPOVICI, Christian. « La chaîne de traitement des documents numériques : Caractéristiques et mise en œuvre ». *Bulletin des bibliothèques de France* [en ligne]. 2002, n° 1, p. 86-91. [Consulté le 14 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

MASANES, Julien. *L'archivage des sites Internet : rapport de stage*. Villeurbanne : Enssib, 1999, 24 p.

NAWROCKI, François. *Le protocole OAI et ses usages en bibliothèque* [en ligne]. Paris : Ministère de la culture et de la communication, 2005. Mise à jour le 15 février 2005. [Consulté le 23 novembre 2005]. Disponible sur : <<http://www.culture.gouv.fr/culture/dll/OAI-PMH.htm>>.

Aspects juridiques :

ALIX, Yves. « La banalisation des bibliothèques : une offre non marchande dans le marché de l'information ». *Bulletin des bibliothèques de France* [en ligne]. 2002, n° 1, p. 23-29. [Consulté le 14 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

ALIX, Yves (dir.). *Le droit d'auteur et les bibliothèques*. Paris : Electre – Cercle de la Librairie, 2000, 237 p.

BATTISTI, Michèle. « Le droit d'auteur, un obstacle à la liberté d'information ? », *Bulletin des bibliothèques de France* [en ligne]. 2004, n°6, p. 31-35. [Consulté le 23 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

BERTRAND, Anne-Marie. « Les lecteurs, le droit et l'argent », *Bulletin des bibliothèques de France* [en ligne]. 2004, n° 6, p. 42-50. [Consulté le 23 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

DANSET, Françoise. « La directive sur le droit d'auteur et les droits voisins : Un combat nécessaire », *Bulletin des bibliothèques de France* [en ligne]. 2004, n° 6, p. 36-41. [Consulté le 23 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

FRANCE. Ministère de la culture et de la communication. *Projet de loi relatif au droit d'auteur et aux droits voisins dans la société de l'information* [en ligne]. [Consulté le 9 décembre 2005]. Disponible sur : <<http://www.culture.gouv.fr/culture/actualites/index-lois.htm>>.

FRANCE. Ministère de la culture et de la communication. *Questions juridiques liées à l'exploitation des documents numériques* [en ligne]. Paris : Ministère de la culture et de la communication, 2005. [Consulté le 23 novembre 2005]. Disponible sur : <<http://www.culture.gouv.fr/culture/mrt/numerisation/fr/dll/juridi.html>>.

MARTER, Alain, SALAÜN, Jean-Michel. « Propriété intellectuelle et bibliothèques françaises : leçons américaines et opportunités européennes ». *Bulletin des bibliothèques de France* [en ligne]. 1998, n° 3, p. 12-16. [Consulté le 14 novembre 2005]. Disponible sur : <<http://bbf.enssib.fr>>.

PARLEMENT EUROPÉEN ET CONSEIL DE L'UNION EUROPÉENNE. *Directive 2001/29/CE du Parlement européen et du Conseil du 22 mai 2001 sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information* [en ligne]. Journal officiel des Communautés européennes, n°L167 du 22 juin 2001, p. 10-19. [Consulté le 14 novembre 2005]. Disponible sur : <http://europa.eu.int/eur-lex/pri/fr/oj/dat/2001/l_167/l_16720010622fr00100019.pdf>.

Table des annexes

ANNEXE 1 : GLOSSAIRE.....	45
ANNEXE 2 : RÉUNIONS EFFECTUÉES PAR LE STAGIAIRE.....	46
ANNEXE 3 : MIGRATION DES DOCUMENTS NUMÉRIQUES SUR LES SERVEURS DU DSI.....	48
ANNEXE 4 : LE MODÈLE OAIS.....	49
ANNEXE 5 : DOCUMENTS DE TRAVAIL SUR LES ARBRES DE DÉCISION. MODE D'EMPLOI.....	50
ANNEXE 6 : DOCUMENTS DE TRAVAIL SUR LES ARBRES DE DÉCISION. NOTE DE CADRAGE DU MODULE DE GESTION DE DROITS	54
ANNEXE 7 : DOCUMENTS DE TRAVAIL SUR LES ARBRES DE DÉCISION. COMPTE RENDU DE LA RÉUNION DU 23 NOVEMBRE.....	58
ANNEXE 8 : DOCUMENTS DE TRAVAIL SUR LES ARBRES DE DÉCISION. LES 6 ARBRES.....	64

Annexe 1 : glossaire

ABN : Agence bibliographique nationale

AFNIC : Association française pour le nommage internet en coopération. Elle est chargée, au niveau français, d'attribuer les noms de domaine en .fr

Crawl : Collecte de sites internet

DBN : département de la bibliothèque numérique

DL : dépôt légal

DCO : Direction des collections

DSR : Direction des services et des réseaux

DSI : département des systèmes d'information

Frame : Système de division d'une page Web en deux ou plusieurs fenêtres indépendantes.

OAI-PMH : Open Archives Initiative – Protocol for Metadata Harvesting.
Protocole d'échange de métadonnées

OAIS : modèle de référence pour les systèmes ouverts d'archivage (en anglais : Reference Model for an Open Archival Information System).

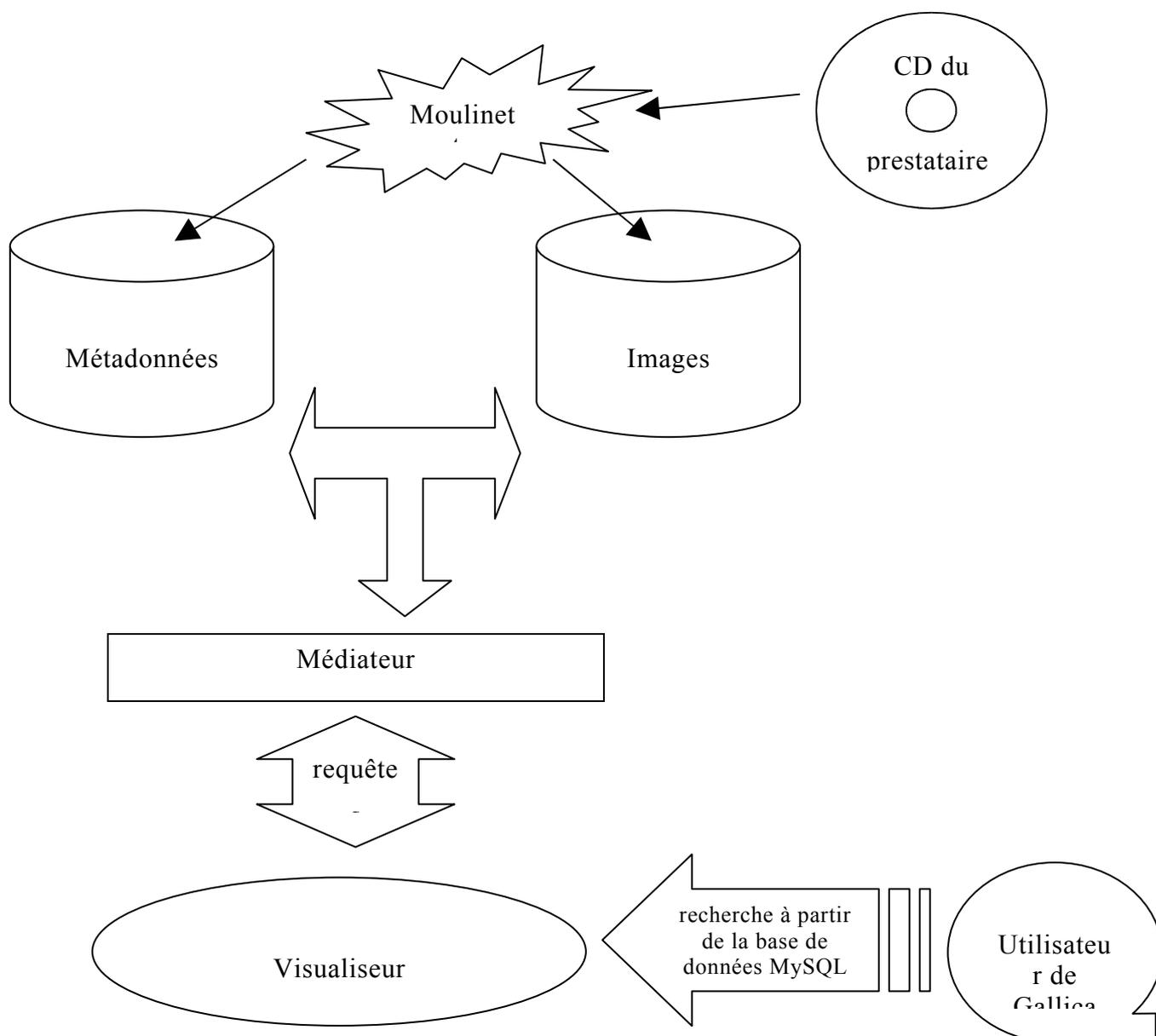
URL : Uniform resource locator. « Adresse » d'une page Web.

Annexe 2 : réunions effectuées par le stagiaire

Vendredi 9 septembre	Réunion de service du DBN.
Mercredi 14 septembre	Chaîne de traitement du document numérique : rencontre avec des responsables des départements de la coopération et de la conservation (Catherine Mocellin, Nathalie Leborgne, Hélène Dontenville). Chaîne de traitement du document numérique : rencontre avec Marie-Élise Fréon.
Jeudi 22 septembre	Intégration de la banque d'images dans l'entrepôt OAI : réunion avec Yves Le Guillou, adjoint au directeur du département de la reproduction.
Lundi 26 septembre	Préparation des arbres de décision : réunion avec Valérie Game et Sophie Sepetjan.
Mardi 27 septembre	Matin. Présentation de la BnF : budget / politique documentaire / acquisitions / documentation électronique. Présentation par François Mathon, Caroline Rives, Marie-Dominique Loret, Coralie Dunan Vauchelles, Olivier Jacquot. Après-midi. Visite du centre de Bussy-Saint-Georges.
Mercredi 28 septembre	Présentation de la BnF : traitement des collections. Présentation par Danièle Heller et Christian Lupovici.
Jeudi 29 septembre	Réunion du groupe de travail OAI-PMH.
Lundi 3 octobre	Présentation de la BnF : numérisation / coopération. Présentation par Claire Simon et Frédérique Joannic-Seta.
Jeudi 6 octobre	Réunion de service du DBN.

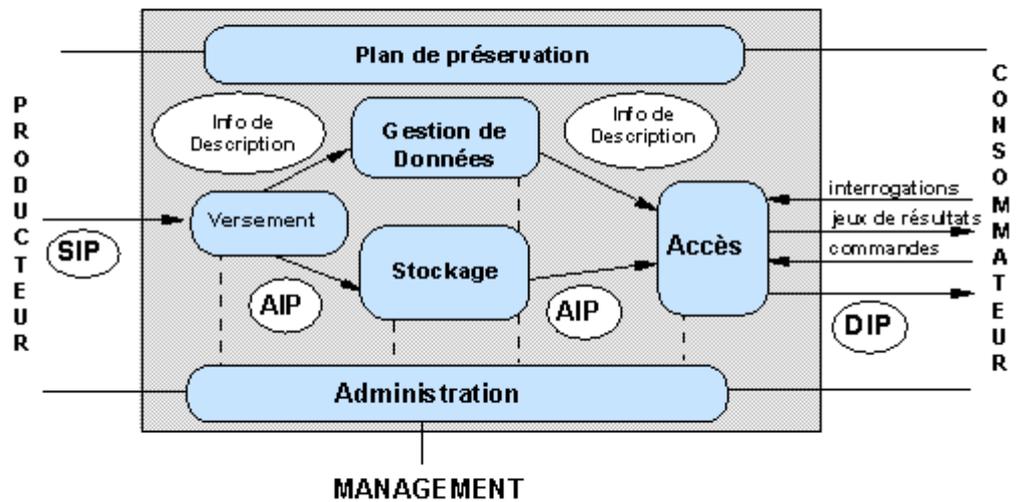
Vendredi 7 octobre	Réunion de service de la Bibliographie française livres – département du dépôt légal.
Lundi 10 octobre	Préparation des arbres de décision : réunion avec Valérie Game et Sophie Sepetjan.
Mardi 11 octobre	Présentation de la BnF : services au public sur place et à distance. Présentation par Lenka Bokova et Isabelle Copin.
Jeudi 13 octobre	Réunion de service du DBN.
Mardi 18 octobre	Présentation de la BnF : politique d'accréditation / étude des publics. Présentation par Claude Collard et Romuald Ripon.
Lundi 24 octobre	Préparation des arbres de décision : réunion avec Valérie Game et Sophie Sepetjan.
Mardi 25 octobre	Visite du département de l'Audiovisuel. Présentation par Isabelle Giannattasio-Mazeaud.
Vendredi 28 octobre	Conférence du C.I.D. sur la gestion des droits des documents numérisés.
Lundi 7 novembre	Préparation des arbres de décision : réunion avec Valérie Game et Sophie Sepetjan.
Vendredi 18 novembre	Harmonisation des déclarations de droits sur le site de la BnF : réunion avec Valérie Game et Sophie Sepetjan.
Mardi 22 novembre	Réunion d'information sur les identifiants pérennes par Emmanuelle Bermès et Laurent Duplouy.
Mercredi 23 novembre	Présentation des arbres de décision lors de la réunion du groupe de travail sur les droits des documents numériques.

Annexe 3 : migration des documents numériques sur les serveurs du DSI



Annexe 4 : le modèle OAIS

Le modèle fonctionnel OAIS



Annexe 5 : Documents de travail sur les arbres de décision. Mode d'emploi

Règles générales

Objectif des arbres de décision

Les arbres de décision sont une forme de modélisation des métadonnées de gestion de droits. L'arbre de décision revient à poser un certain nombre de questions sur un document, dans un certain ordre, de façon à déterminer sa communicabilité.

Pour obtenir la limite de communication la plus favorable possible, on a cherché à prévoir le plus grand nombre de cas de figure.

Les arbres de décision doivent être disponibles pour tous les types de documents :

- imprimés
- manuscrits
- photographies non publiées et publiées
- documents sonores inédits et édités
- images animées inédites et éditées
- œuvres en trois dimensions
- archives de la Bibliothèque nationale de France

Les partitions sont soit imprimées, soit manuscrites.

Les œuvres en trois dimensions comprennent aussi bien les monnaies et médailles que les objets divers contenus dans des fonds de papiers d'écrivains (masques mortuaires...).

Il a été décidé de séparer les cas de figures édités / non-édités, car cette caractéristique génère des problèmes spécifiques. Les documents publiés ou édités ont des périodes de protection parfois plus longues (cas des prorogations de guerre), mais les documents non-édités posent plus de problèmes de droit à l'image ou au respect de la vie privée.

Pour l'instant, seuls les arbres des imprimés, des manuscrits, des photographies non publiées et publiées, des documents sonores édités et des images animées éditées, ont été établis.

Respect des différents types de droit

Les différents types de droit ont été pris en compte :

- Droit patrimonial (dépendant du droit d'auteur) : toute œuvre tombe dans le domaine public 70 ans après la mort de son dernier auteur, mais il existe des prorogations de guerre (dans le cas où l'auteur est mort pour la France, ou dans le cas où le livre a été publié avant une guerre mondiale, car les périodes de guerre ne permettent pas une exploitation correcte des œuvres de l'esprit).
- Il ne faut pas négliger que le droit patrimonial s'attache aussi bien à l'œuvre elle-même (par exemple, une photographie que l'on va numériser) qu'à toute œuvre qui y serait reproduite (un tableau reproduit sur une photographie).
- Droit à l'image et au respect de la vie privée. Il s'éteint à la disparition de la personne. Mais il y a des exceptions : le traitement inhumain ou dégradant, l'atteinte à la réputation d'une personne. Dans ce cas, il est nécessaire de prévoir une date de sécurité très longue : 270 ans (ce qui donne, pour les photographies et les images animées, qui sont nécessairement moins vieilles : communication soumise à autorisation).
- Droits voisins, reconnus à certains auxiliaires de la création littéraire et artistique : artistes-interprètes, producteurs de phonogrammes ou de vidéogrammes, entreprises de communication audiovisuelle.

- Décisions de justice : il est possible que par une décision de justice, certains ouvrages aient été interdits (incitation à la pédophilie, à la haine raciale, atteinte à la vie privée...). Dans ce cas, ces ouvrages sont inaccessibles.

Règles particulières

Présence de conditions écrites

Différentes conditions écrites peuvent modifier le caractère de communicabilité des documents. Il peut s'agir de conditions spécifiques pour le dépôt d'un manuscrit, de stipulations lors d'un legs, ou encore d'un accord particulier avec un éditeur.

Les conditions écrites peuvent lever des interdictions de communication relevant de différents types de droits, ou au contraire bloquer des œuvres plus longtemps que ce que les règles du droit d'auteur prévoiraient. C'est pourquoi il a été décidé de les traiter parallèlement aux règles générales.

La date de communicabilité prévue par ces conditions écrites sera calculée et inscrite « manuellement ».

Clauses d'invalidation

Pour choisir, le système de gestion des droits doit avoir des priorités.

- Les conditions écrites sont prioritaires par rapport aux règles générales, que ces conditions donnent une date de communication plus précoce ou plus tardive.
- Entre deux dates de communication établies grâce à la règle générale, on choisit la plus restrictive. Ceci peut arriver dans quelques cas assez rares. Par exemple, imaginons le cas d'un auteur mort pour la France depuis plus de 100 ans, dont l'œuvre a fait l'objet d'une édition scientifique en 1960. En suivant le chemin « auteur mort pour la France / mort pour la France depuis plus de 100 ans », on obtient la réponse : « document communicable Rech/BnF/Net ». Toutefois, le chemin « publié après 1948 / dernier auteur mort depuis moins de 70 ans » conduira à : « Non communicable ».
- Il existe, pour certaines décisions, une clause « d'invalidation », qui est prioritaire sur les éventuelles « conditions particulières ». En effet, même dans

le cas où l'on dispose d'autorisation de communication de documents, il n'est pas toujours possible d'être sûr que ceux qui les ont donné ont bien *tous* les droits sur ces mêmes documents. Par exemple, un photographe peut céder son fonds de clichés, en autorisant la BnF à les mettre en ligne. Mais les personnes qui peuvent avoir été photographiées n'ont probablement pas donné leur accord pour que leur image soit mise sur Internet. C'est pourquoi avant une date de sécurité de 80 ans (puisque le droit à l'image s'éteint à la mort de la personne), il sera impossible de mettre sur Internet des images représentant des personnes.

Enfin, il faut souligner que dans le domaine juridique, l'automatisation complète des décisions est impossible. C'est pourquoi il sera possible à l'administrateur du système de le contraindre « manuellement », au cas par cas, d'interdire ou d'autoriser une communication.

Légende des arbres :

DL : dépôt légal.

Communicable Rech : communicable en Rez-de-Jardin et dans les salles de lecture des départements spécialisés de Richelieu.

Communicable BnF : communicable en Haut et Rez-de-Jardin.

Communicable Net : mis en ligne sur Internet.

Annexe 6 : Documents de travail sur les arbres de décision. Note de cadrage du module de gestion de droits

DSR/DBN/Emmanuelle Bermès

DOCUMENT DE TRAVAIL

Note de cadrage du module de gestion de droits

Le système de gestion de droits (ci-après nommé G2D) préparé par le groupe est un **module** du système de préservation à long terme des documents numériques.

Il constituera le **référentiel unique** de gestion des droits d'accès et d'usage des documents numériques dans le système.

Métadonnées de gestion des droits

Pour favoriser son autonomie et la cohérence des données, G2D disposera de sa propre **base de métadonnées**, dans laquelle seront stockées pour chaque document les métadonnées utiles pour déterminer sa communicabilité.

Ces métadonnées seront

- extraites automatiquement d'autres sources au moment du versement (catalogue, métadonnées de préservation)
- renseignées manuellement dans G2D grâce à des interfaces spécifiques.

Règles de décision

Les différentes étapes nécessaires pour déterminer la communicabilité d'un document peuvent être modélisées sous forme **d'arbres de décision**.

Les arbres de décision seront créés par type de document ; il y en aura une dizaine au total.

Contraintes (droits d'usages)

Les contraintes sont des limitations d'usage appliquées une fois que l'accès au document a été autorisé.

Ces contraintes peuvent porter sur l'impression, le téléchargement, la résolution, le prix...

Il y a deux types de contraintes :

- **contraintes par défaut** : ce sont des simples paramètres de communication s'appliquant à tous les types de documents si aucune précision complémentaire n'est disponible
- **contraintes spécifiques** : liées à un projet, un contrat...

Ensembles

G2D permettra de regrouper des **ensembles de documents** pour leur appliquer des conditions d'accès communes en plus des règles générales (par exemples dans le cas d'un contrat).

Les ensembles auront leurs propres métadonnées décrites à part.

Contexte d'accès

Le contexte d'accès correspond à l'addition de deux critères :

- le lieu où se trouve le demandeur : **profil d'accès**
- le niveau d'accréditation du demandeur : **profil utilisateur**

Fonctionnement du service de G2D

Le service que G2D fournit est un **service d'information** à destination des autres modules du système de préservation et en particulier les modules d'accès. Ce

service consiste à communiquer à l'application qui lui en fait la demande toutes les conditions possibles de communication d'un document dans différents contextes et suivant différentes contraintes.

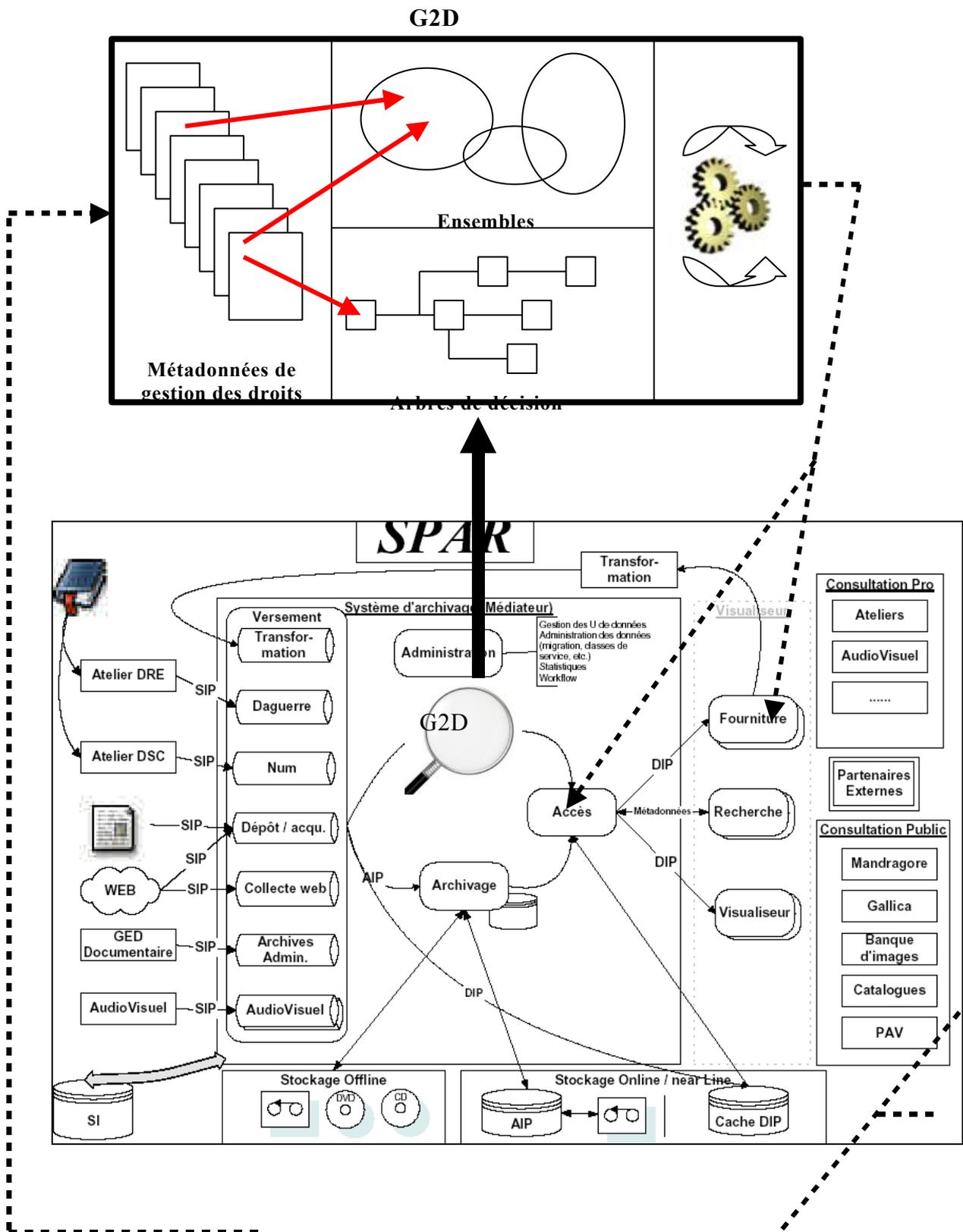
Le service se déroule de la manière suivante :

- l'application d'accès (catalogue, médiateur ou autre) envoie l'identifiant d'un document (ou d'une portion de document) à G2D
- G2D utilise les métadonnées du document, l'arbre de décision concerné, éventuellement les ensembles auxquels appartient le document pour déterminer dans quel(s) contexte(s) et suivant quelle(s) contrainte(s) le document est consultable
- Il transmet ces informations à l'application d'accès
- L'application d'accès **doit prendre en charge** la détection du contexte d'accès, la limitation des accès suivant les contextes et les contraintes, l'affichage des messages d'information et d'erreur.

Administration

G2D fournira des services d'administration tels que

- niveaux d'utilisateurs administrateurs / gestionnaires / correspondants juridiques (à définir sur le plan organisationnel)
- programmation d'alertes
- interfaces de saisie et de correction de métadonnées
- service de communication exceptionnelle ?
- mode de fonctionnement dégradé en cas d'accident du système.



Annexe 7 : Documents de travail sur les arbres de décision. Compte rendu de la réunion du 23 novembre

C Bibliothèque nationale de France
direction des services et des réseaux
département de la bibliothèque numérique

quai François Mauriac 75706 Paris
cedex 13
téléphone 33 (0) 1 53 79
télécopie 33 (0) 1 53 79 TAD

mél frederic-
d.martin@bnf.fr

Réf. :	Affaire suivie par : Frédéric Martin, Emmanuelle Bermès Tél : 01 53 79 49 77	compte rendu de réunion
Paris,	le 5/12/05	
réunion	Groupe de travail « Gestion des droits des documents numérisés » 23/11/05	
participants	Emmanuelle Bermès (DBN) Marie-Elise Fréon (DSC/NUM) Michèle Guy (ABN) Frédérique Laval (DCO/Projet Richelieu) Frédéric Martin (DBN) Clément Oury (conservateur stagiaire)	Laurent Duploux (DSI) Valérie Game (SJ) Catherine Lupovici (DBN) Dominique Maillet (DSC/SEC) Marie-Claude Thompson (DCO/EST)
copies pour information		
prochaine réunion	à fixer selon les besoins	

Ordre du jour

La seconde réunion du groupe de travail « Gestion des droits des documents numérisés » avait pour objet :

- la présentation générale du futur système de gestion des droits numérisés (G2D)
- l'examen de six arbres de décision permettant de déterminer le statut de communicabilité des documents
- les prochaines étapes du projet et les actions à mener

Synthèse des discussions

La réunion débute par un bilan d'avancement du projet depuis la dernière réunion, qui s'est tenue le 24/06/05. Trois types d'actions ont été menées parallèlement : la conceptualisation et la modélisation générale de G2D (voir 2.1), l'élaboration d'un premier jeu de six arbres de décisions (voir 2.2), l'harmonisation des déclarations de droits actuelles sur les différents sites et sous-sites internet de la BnF.

Le DBN et le Service juridique ont initié une réflexion sur la mise en cohérence des formulations juridiques existantes. Ce travail est nécessaire pour rendre plus claire la distinction entre plusieurs niveaux de protection : les droits sur le site lui-même en tant qu'œuvre de l'esprit (charte graphique, contenus éditoriaux) ; sur le site en tant que base de données ; sur les ressources qu'il rend accessibles. Par ailleurs, l'harmonisation des mentions existantes est une étape préalable à la création de nouvelles déclarations de droits, par exemple pour les entrepôts de métadonnées OAI-PMH ou dans le système G2D lui-même. Une proposition sera soumise au groupe de travail, probablement lors de la prochaine réunion.

Le DBN tient à souligner la qualité et le sérieux du travail effectué par Clément Oury, conservateur en stage au DBN depuis le début du mois de septembre, qui a participé à la création des arbres de décision. Cette réunion est donc l'occasion pour lui de présenter son travail, avant son retour prochain à l'ENSSIB.

Modélisation du système G2D

E. Bermès (DBN) présente le résultat du travail mené avec L. Duplouy (DSI) concernant la modélisation du système G2D. Une note de cadrage est distribuée à ce sujet aux participants. Elle est également fournie en pièce jointe au présent compte-rendu.

G2D sera le **référentiel unique** de gestion des droits d'accès et d'usage des documents numériques, dans le futur système de stockage pérenne et d'accès aux documents numériques, dont il constituera un module. Il possédera **sa propre base de métadonnées**, déterminera les statuts de communicabilité en fonction de règles de décision modélisées par des **arbres de décisions**, et pourra intégrer des contraintes d'usage (impression, téléchargement, résolution...).

Il a paru important de prévoir la **gestion d'ensembles documentaires**, pour lesquels des conditions juridiques particulières doivent s'appliquer (contrats). Des règles seront définies et stockées pour chacun de ces ensembles, qui prendront le pas sur les règles générales. Pour l'instant, la présence possible de « conditions écrites » est prévue par les arbres de décisions ; il paraît plus pertinent de les en dissocier.

GD2 fournira un **service d'information** à l'application d'accès. Il faudra définir différents profils d'utilisateurs et d'administrateurs pour la gestion de ce système.

Discussion

M.-E. Fréon (DCP/NUM) pose la question des métadonnées de droits déjà définies dans les métadonnées de préservation et de leur utilisation / réutilisation au sein de G2D. Il est envisagé de revoir ces métadonnées en prenant en compte les résultats du projet G2D.

D. Maillet (DSC/SEC) demande s'il sera possible d'interroger le système pour connaître les droits des documents avant même leur numérisation. Cette utilisation permettrait de savoir *a priori* si ces documents sont susceptibles d'avoir les statuts de communicabilité les plus intéressants pour les lecteurs ; elle fournirait une aide supplémentaire pour définir les priorités dans les programmes de numérisation. Cette possibilité sera étudiée et éventuellement intégrée dans les spécifications fonctionnelles.

Examen des arbres de décision

La seconde partie de la réunion est consacrée à la discussion des arbres de décision, qui ont été élaborés par le Service juridique (V. Game et S. Sepetjan) et le DNB (C. Oury et F. Martin). Ces arbres constituent une première représentation graphique de l'ensemble des données nécessaires à la détermination du statut de communicabilité par le système et de leur enchaînement. Ils ne seront donc pas traités tels qu'elles par le système, mais devront faire l'objet d'une formalisation informatique et logique. Ce formalisme précisera notamment quand les cheminements sont alternatifs (exclusifs les uns des autres) et quand ils doivent être parcourus en parallèle afin d'aboutir à trois réponses, une par type d'accès (recherche / BnF / Internet).

Six arbres de décision ont été réalisés jusqu'à présent :

- les imprimés

- les manuscrits
- les photographies publiées
- les photographies non publiées
- les images animées éditées
- les images animées non éditées

Pour en illustrer le fonctionnement, Clément Oury présente quelques exemples :

- l'édition de 1853 de *Le Rouge et le Noir* de Stendhal ; d'après l'arbre des imprimés, la version numérique de ce document est accessible via les trois types d'accès : recherche / BnF / Internet.
- l'édition de 2001 du même ouvrage par Maxi-Livres est également disponible par ces trois types d'accès, car Maxi-Livres ne réédite que le texte brut des grands classiques, sans appareil critique, préface ni postface (donc sans auteurs autres que l'auteur principal)
- l'arbre des photographies est le plus complexe, car il doit prendre en compte différents types de droits : les droits d'auteurs du photographe, le cas échéant le droit à l'image des personnes photographiées, le cas échéant les droits d'auteur(s) des œuvres de l'esprit photographiées. Ce type de documents rend également nécessaire le recours aux « clauses d'invalidation », qui s'ajoutent aux règles générales et aux éventuelles conditions écrites.

En effet, les conditions écrites fixées par les contrats ne prévoient pas toujours les questions liées au droit à l'image et aux droits associés aux œuvres de l'esprit photographiées. Pour éviter une mise à disposition automatique de contenus problématiques, les clauses d'invalidation constituent une sécurité supplémentaire qui peut être levée par les gestionnaires du système, au cas par cas, lorsque cela se justifie.

Par exemple, le recueil de photographies *Epouvantails* de Robert Doisneau ne pose pas de problèmes particuliers quant au droit à l'image. Il en va autrement du recueil *Artistes dans leur atelier*, du même photographe, qui donne à voir des photographies de personnes identifiées et de leurs propres œuvres.

Les arbres de décisions présentés sont validés par le groupe de travail. Il est décidé d'ajouter une fonction du type « communication exceptionnelle », si elle est techniquement réalisable. Elle permettrait notamment l'accès ponctuel, sur demande motivée, à des ressources pour lesquelles la durée de sécurité est particulièrement longue (par ex., les documents susceptibles de porter atteinte à l'image ou à la réputation des personnes, en principe inaccessibles pendant 270 ans).

Discussion

Concernant les estampes, la question est de savoir s'il faudra les traiter plutôt suivant l'arbre des imprimés ou celui des photographies. On peut supposer que le droit à l'image ne s'applique pas aux personnes figurant sur les estampes, où la démarche artistique l'emporte sur la représentation

« réaliste ». Marie-Claude Thompson (DCO/EST) pose également la question des affiches. De manière générale, au moment d'aborder les règles de décision portant sur les documents des départements spécialisés, il sera nécessaire de solliciter l'expertise des chargés de collections au sein de ces départements. Il en va de même pour les documents numériques natifs, dont les règles de décision seront fixées en collaboration avec le Département de l'Audiovisuel.

Il ne faudra pas oublier non plus les documents administratifs produits par les agents de la BnF, qui seront pour certains versés dans le futur système de stockage pérenne. Un arbre spécifique sera donc conçu pour ces documents, avec l'aide de Catherine Dhérent (DGE/Mission ARC).

L'étape ultérieure consistera, à partir des données apparaissant dans les arbres de décision, à lister les métadonnées de droit nécessaires à chaque type de document, en distinguant :

- les informations communes à tous les documents et celles qui varient d'un arbre à l'autre
- les informations qui peuvent être récupérées d'une source externe et celles qui doivent être saisies manuellement
- les informations obligatoires pour le versement dans SPAR et celles qui sont facultatives

Cette étape est nécessaire pour concevoir les formulaires de saisie et de visualisation et instruire la question du formalisme XML : adoption/adaptation/création d'un *Rights Expression Language (REL)*.

Une fonction d'alerte pourrait être prévue pour le cas où G2D ne disposerait pas d'informations suffisantes, au moment de la communication. Les gestionnaires pourraient ainsi identifier *a posteriori* des lacunes à combler, concernant des documents ayant fait l'objet de demandes d'accès. Cette fonctionnalité implique de savoir si les métadonnées manquantes peuvent être renseignées (moyennant des recherches plus approfondies) ou non.

Liste des actions décisions

Les actions à faire (introduites ce jour) sont représentées par le symbole

Les actions terminées sont représentées par le symbole

N°	Nature (action ou décision)	Qui	Quand	Fait
1	Rédaction des spécifications fonctionnelles de G2D	DBN/DSI		<input type="checkbox"/>
2	- Arbre de décision des documents audiovisuels - Arbre des documents sonores	SJ/DBN/DAV		<input type="checkbox"/>
3	Arbre de décision des archives administratives	SJ/DBN/ARCH		<input type="checkbox"/>
4	- Arbre des objets en 3D (monnaies, médailles...)	SJ/DBN/Projet Richelieu + experts		<input type="checkbox"/>
5	Rédaction de consignes pour l'harmonisation des déclarations de droit	SJ/DBN		<input type="checkbox"/>
6	Liste des métadonnées de droit nécessaires à chaque type de document	DBN/DSC		<input type="checkbox"/>
7	Etude des principaux REL existants	DBN		<input type="checkbox"/>

Annexe 8 : Documents de travail sur les arbres de décision. Les 6 arbres

Naturels

