



COMMISSION BIBLIOTHEQUES NUMERIQUES RAPPORT 2011

Introduction

Le rapport « *Schéma numérique des bibliothèques* » a été présenté par Monsieur Bruno Racine, président de la Bibliothèque nationale de France, au Conseil du Livre le 22 mars 2010. Il avait été élaboré par un groupe interministériel *ad hoc* associant une vingtaine de représentants du secteur des bibliothèques et du livre, dont des représentants du Ministère de la Culture et de la Communication et du Ministère de l'Enseignement Supérieur et de la Recherche. Ce rapport dresse un état des lieux des collections et des usages numériques en bibliothèque, et émet des recommandations notamment dans le domaine de l'acquisition des ressources électroniques, de la conservation numérique et d'une meilleure concertation en matière de numérisation.

Les recommandations du Schéma numérique des bibliothèques constituent un socle solide pour le *Plan pour le renouveau des bibliothèques* rendu public en février 2010 par le MESR, qui comporte un engagement fort en faveur du développement massif de l'offre numérique, comme pour les *14 propositions pour le développement de la lecture* présentées par le MCC le 30 mars 2010, qui proposent aux collectivités territoriales un contrat numérique pour les bibliothèques et envisagent de développer et coordonner les services numériques.

Dans ce cadre, les deux ministères ont décidé de travailler en étroite concertation pour étudier ensemble les suites données aux recommandations du rapport Racine et pour déterminer les cadres d'action commune, offrir des instruments collectifs, et mutualiser les efforts et les coûts. Est ainsi mise en œuvre la première recommandation du rapport, qui préconisait de « *pérenniser la concertation autour du numérique à travers un groupe de contact permanent sur les bibliothèques numériques*. ».

Suite à un échange de courriers entre les deux ministres en juin-juillet 2010, les deux directeurs de cabinet ont confié par lettre de mission le 8 novembre 2010 à Monsieur Pierre Carbone, inspecteur général des bibliothèques, « *le soin de présider une commission « Bibliothèques numériques » composée à parité de représentants des deux ministères et chargée d'étudier les suites données aux 11 recommandations formulées dans le rapport*. »

La commission réunit les membres suivants :

Pierre CARBONE, président (IGB) ; Nicolas GEORGES (DGMIC-SLL) ; Michel MARIAN (MISTRD) ; Francis ANDRE, (DIST-CNRS) ; Patrick BAZIN (BPI) ; Marine BEDEL (Ville de Rennes) ; Raymond BERARD (ABES) ; Catherine DHERENT (BnF) ; Jacques FONTANILLE (CPU) ; Marie-Dominique HEUSSE (PRES Université de Toulouse) ; Valérie TESNIERE (BDIC) ; Maïté VANMARQUE (Ville de Rouen) ; Mélanie VILLENET-HAMEL (Département de l'Hérault).

Le secrétariat est assuré par Fabien PLAZANNET (DGMIC-SLL) et Alain COLAS (MISTRD).

La commission a été installée le 15 février 2011 et a tenu ensuite quatre séances de travail : le 30 juin 2011 sur l'archivage pérenne des documents numériques, le 12 octobre 2011 sur l'acquisition des ressources électroniques, le 1^{er} décembre 2011 sur la numérisation, le 2 février 2012 sur le projet de rapport annuel.

Le rapport 2011 présente un récapitulatif des débats ainsi que des propositions d'action. Certaines de ces propositions constituent des orientations pour la poursuite des travaux de la Commission en 2012, d'autres visent à s'appuyer sur la Commission pour le lancement d'actions qui seraient mises en œuvre par les acteurs concernés. Il comprend également en annexe une synthèse des auditions.

Récapitulatif des débats et propositions de la commission

1. L'archivage pérenne des documents numériques

Les documents produits ou acquis sous forme numérique connaissent une croissance exponentielle, et la garantie de leur archivage pérenne est d'importance vitale pour la préservation et la transmission du patrimoine des bibliothèques. Il en est de même pour les copies numériques de documents sur support matériel, pour lesquels la conservation des exemplaires originaux reste encore nécessaire. Les moyens humains engagés par les opérateurs les plus importants (des équipes d'une dizaine d'ETP) ainsi que les coûts d'investissement et de fonctionnement justifient de renforcer les démarches de mutualisation.

La mutualisation de l'archivage comme sa gestion sur la longue durée sont indispensables.

Plusieurs acteurs publics nationaux sont identifiés en matière d'archivage pérenne ou intermédiaire.

Le **Service Interministériel des Archives de France (SIAF)** est en charge de la politique nationale des archives (hors celles de la Défense et des Affaires étrangères), audite les systèmes d'archivage et donne son agrément aux tiers-archivistes publics ou privés pour la conservation d'archives publiques courantes et intermédiaires.

Parmi les **tiers-archivistes**, les opérateurs publics ont leur place, notamment pour les bibliothèques le Centre informatique national de l'enseignement supérieur – CINES- (qui a reçu l'agrément SIAF pour la conservation d'archives publiques intermédiaires) et la Bibliothèque nationale de France – BnF – (qui va demander l'agrément SIAF).

D'autres acteurs publics assurent un archivage intermédiaire ou pérenne, notamment : le Centre national d'études spatiales (CNES) avec sa plateforme STAF en production depuis 1995, l'Institut national de physique nucléaire et de physique des particules (IN2P3) au Centre national de la recherche scientifique (CNRS), la Direction générale des finances publiques (DGFIP) avec sa plateforme ATLAS lancée en 2007, l'Imprimerie nationale, l'Institut géographique national (IGN).

D'autres structures publiques jouent un **rôle d'agrégateur** entre ces opérateurs d'archivage et les producteurs de données, notamment dans le domaine de la recherche, telles le Centre de communication scientifique directe du CNRS ou le Très grand équipement (TGE) ADONIS.

Au niveau des **bibliothèques**, il existe des solutions publiques de tiers-archivage pour les bibliothèques de l'enseignement supérieur et la recherche, ainsi que pour les bibliothèques menant des opérations en partenariat avec la BnF. Mais nombre de bibliothèques territoriales ou universitaires engagées dans des opérations importantes de numérisation ne disposent pas actuellement de telles solutions, et assurent un simple stockage de données dont le volume est parfois important.

La cartographie des compétences élaborée pour le Schéma numérique des bibliothèques doit être mise à jour et complétée en y intégrant les autres acteurs publics (CNES, IGN, etc....) de façon à disposer d'un état des lieux plus complet sur les masses de données archivées, les critères mis en œuvre et les coûts. Cette cartographie des compétences pourra définir des règles concernant les responsabilités de tiers-archivage. Pour éviter la dispersion des efforts, cette offre de tiers-archivage se ferait selon une répartition des responsabilités et des communautés concernées, en identifiant des opérateurs par communauté pour sécuriser les financements, selon les préconisations du Schéma numérique des bibliothèques et du projet de la Bibliothèque scientifique numérique (BSN).

Assurer la viabilité de l'offre de tiers-archivage par les opérateurs publics nécessite de travailler aux modèles économiques qu'ils proposent, d'analyser les coûts complets et les modes d'amortissement des investissements du côté du prestataire, mais aussi d'analyser les tarifs et leur soutenabilité à long terme pour les utilisateurs, en tenant compte des offres proposées par d'autres prestataires publics ou privés en France ou à l'étranger.

Proposition 1 : Mener en 2012 une étude des modèles économiques de tiers-archivage public.

La Commission Bibliothèques numériques envisage de mener cette étude dans le cadre de la poursuite de ses travaux

A la suite du rapport de Maurice Quenet « Quel avenir pour les Archives de France ? », le nouveau chantier que constitue l'archivage numérique définitif exige un nouveau mode de gouvernance. Si la diversité des architectures techniques est réelle, il serait justifié de **créer un centre de compétences mutualisé pour la préservation de l'information numérique**, réunissant les acteurs institutionnels concernés. Un tel centre serait compétent pour la veille sur les formats d'archivage, les supports, et les formats de métadonnées et pour la formation.

La mise en place d'une structure fédérant les acteurs concernés est indispensable à une coopération stratégique au niveau national. Cette mise en œuvre pourrait être organisée par la DISIC, Direction interministérielle des systèmes d'information et de communication, créée par décret du 21 février 2011. Celle-ci est chargée notamment d'organiser toutes les opérations de mutualisation des systèmes d'information entre les différentes administrations de l'Etat. Elle pourrait s'appuyer pour ce centre de compétence sur la BSN, sur le CINES, principal opérateur d'archivage pérenne de l'enseignement supérieur et de la recherche, sur le SIAF et sur la BnF pour les établissements culturels. Ce centre serait appelé à assurer des missions sur le long terme, en accord avec la temporalité longue des archives, et à favoriser des recherches pour faire face aux doutes technologiques.

Deux niveaux de mutualisation peuvent être identifiés :

- le premier, commun à tous les opérateurs concernés, gérant des missions de mutualisation des compétences : veille, formation, études et recherche, publication de guides de bonnes pratiques, soutien à des programmes de recherches sur les supports, expertise française au niveau européen et international ;
- le second constituant un niveau de coordination et de maîtrise d'ouvrage pour des opérations spécifiques totalement partagées entre plusieurs de ces opérateurs, notamment la coordination de programmes d'archives numériques ou la mutualisation de programmes de numérisation qui prendraient en compte d'emblée la dimension de l'archivage pérenne.

La structure envisagée assurerait la mutualisation de premier niveau et pourrait servir de facilitateur pour la mutualisation de deuxième niveau.

Pour **préfigurer ce centre de compétences**, la première démarche pourrait être de s'appuyer sur un organe réunissant les experts en la matière : le groupe Pérennisation des informations numériques (PIN) au sein de l'association Aristote. Ce groupe, au caractère transverse, réunit des institutions patrimoniales, des organismes à caractère scientifique et technique, de grands groupes nationaux publics ou privés, des experts et chercheurs indépendants reconnus du domaine. Ce groupe métier, qui rassemble des ingénieurs ainsi que des conservateurs des archives et des bibliothèques et des experts de divers ministères, a travaillé sur la norme OAIS et constituerait valablement pour l'Etat un référent pouvant maintenir la cartographie des compétences. Il pourrait être officialisé en qualité de cellule transversale aux ministères et transdisciplinaire, et être missionné dans le cadre de conventions pour des commandes de travaux d'expertise, de veille, de formation, de normalisation en ce domaine, notamment pour des commandes d'études, de formation et de publication conjointes MESR/MCC.

Proposition 2 : Préfigurer un centre de compétences mutualisé en s'appuyant sur le groupe Pérennisation de l'information numérique (PIN).

La Commission Bibliothèques numériques constituera en 2012 en son sein un groupe de travail réunissant les représentants du MCC, du MESR, de la BnF et du CNRS sur ce sujet et prendra l'attache du SIAF. Il s'agira d'en articuler les travaux avec ceux de la DISIC.

2. L'acquisition de ressources électroniques

Pour les bibliothèques de l'enseignement supérieur et de la recherche, la situation évolue fortement, le MESR ayant engagé en 2011 un programme de licences nationales qui va prendre toute son ampleur en 2012-2014 avec les achats prévus dans le programme ISTEX intégré à la BSN, notamment pour les archives de revues et les corpus que seules quelques bibliothèques étaient en mesure d'acquérir. Le dispositif de sélection des licences nationales associe le MCC et la BnF. La sélection des ressources négociées en licence nationale respecte les grands équilibres entre sciences dures et sciences humaines et sociales, entre ressources en anglais et en français. Elles ne concernent pas seulement les grands éditeurs, elles visent aussi à préserver des ressources de niche. L'ensemble des ressources concernées sera valorisé au moyen d'une plateforme nationale à laquelle les portails locaux ou régionaux pourront s'articuler.

Pour les ressources spécifiques aux bibliothèques territoriales, le modèle des licences nationales demande des adaptations. Les bibliothèques publiques sont bien plus nombreuses que les bibliothèques universitaires, leurs usagers sont plus hétérogènes et ont des besoins plus difficiles à cerner. Se pose également la question de la dimension territoriale, les collectivités pouvant mener des politiques documentaires différentes. Par ailleurs, le monde de l'édition générale est plus divers que celui de l'édition scientifique, où dominant quelques oligopoles, les bibliothèques représentent une faible partie du chiffre d'affaires, et le niveau d'implication des éditeurs dans le numérique est très varié. De plus, les bibliothèques publiques ont des attentes spécifiques, portant sur les documents d'auto-formation, l'offre multimédia (VOD et musique en ligne), les jeux et ressources interactive. Les expérimentations actuellement en cours dans un certain nombre de bibliothèques municipales ou départementales de prêt peuvent donner des pistes d'action.

Proposition 3 : Faire le bilan des expérimentations d'achat de ressources électroniques accessibles en ligne aux publics inscrits dans les bibliothèques territoriales

La Commission Bibliothèques numériques envisage de travailler à ce bilan en 2012 dans le cadre de la poursuite de ses travaux.

Dans le programme des licences nationales négociées par le MESR, plusieurs points d'intersection peuvent cependant être trouvés entre bibliothèques publiques et universitaires : certaines bibliothèques municipales bénéficieront ainsi des accords passés avec Springer, ainsi qu'avec Classiques Garnier Numérique. Certaines ressources intéressent à la fois les bibliothèques universitaires et les bibliothèques publiques (par exemple, certains corpus francophones, la presse d'information politique et générale). Le MESR pratique également une politique de soutien aux éditeurs français pour qu'ils proposent une offre de ressources numériques pour les étudiants, en lien avec la mise en place de l'Observatoire des usages du numérique à l'Enssib. Les licences nationales amènent aussi à s'intéresser à l'articulation des ressources de la BnF avec celles des bibliothèques du réseau (notamment les Pôles associés et les CADIST). Mais la BnF ne dispose pas encore des dispositifs techniques qui lui

permettraient de servir à distance ses lecteurs avec les ressources qu'elle a acquises.

Pour répondre à la diversité des besoins, il pourrait y avoir plusieurs types de licences nationales, portées selon les cas par un opérateur unique, par plusieurs opérateurs ou par des groupements d'achats. Mais les licences nationales ne constituent pas la solution unique, et des licences territoriales pourraient être adaptées à des offres de niche (par exemple, la presse quotidienne régionale), à des offres diffusées sur un territoire par des libraires, à la valorisation de contenus locaux. Par ailleurs, les premiers achats en licence nationale concernent des archives ou des corpus constitués, et les attentes des bibliothèques publiques et universitaires se portent aussi sur les ressources courantes.

La réflexion sur les licences nationales ainsi que sur l'accès et la diffusion des ressources numériques en bibliothèque territoriales doit être poursuivie, en donnant plus de place aux collectivités territoriales concernant les ressources présentant un intérêt pour la lecture publique, de façon à formaliser des cas concrets et des modes de négociation.

La question essentielle est en fait celle des usages des publics. Le numérique transforme la lecture publique et les bibliothèques, l'accès d'individus à des capsules de textes change la relation à la communauté de la connaissance. L'appropriation des contenus demande d'autres dispositifs physiques ou numériques, et conduit à mettre l'accent sur l'indexation et la qualité des métadonnées. L'organisation des usages doit faciliter les appropriations et réappropriations de contenus numériques dans un cadre public. Les bibliothèques publiques gardent ainsi un rôle spécifique intellectuel et culturel, par la mise en relation numérique avec des contenus. Il faut donc réfléchir aux usages spécifiques en bibliothèque à l'époque du numérique, ce qui permettra de mieux faire comprendre aux éditeurs et fournisseurs les besoins et attentes en matière d'offre de contenus et de services. La connaissance des usages réels (quelles ressources ? quelles communautés ? quels périmètres d'usage ?) est encore insuffisante. De quoi les usagers ont-ils besoin ? Que font-ils réellement ? Que pourraient-ils faire ? Il serait pertinent d'établir une typologie des usages pour élaborer une stratégie numérique et dialoguer sur cette base avec les éditeurs.

De telles études sont en cours pour la lecture étudiante, notamment à l'Observatoire du numérique dans l'enseignement supérieur créé en 2010 dans le cadre de la plateforme de collaboration entre le Syndicat national de l'édition et le MESR, dont un des objets est de donner la vision d'un manuel numérique. De même, dans le cadre des Investissements d'avenir, est en projet une étude des usages des étudiants dans des sous-domaines (telles les relations internationales) afin de structurer l'offre éditoriale et son mode de présentation.

Il conviendrait également d'étudier les pratiques des usagers non seulement comme consommateurs mais aussi comme producteurs, aussi bien pour ce qui est des étudiants et des chercheurs dans le secteur de l'enseignement supérieur et de la recherche, qu'en lecture publique où les blogueurs et utilisateurs de réseaux sociaux participent à une œuvre collective. Ce rôle de producteurs de contenus est actuellement insuffisamment perçu par les éditeurs comme par les bibliothèques. Des études sont menées localement, par exemple à la BPI ou la BnF ou dans les universités, faisant appel à des focus groupes. Les bibliothèques ont toutefois besoin d'une mutualisation des expériences d'observation des usages, qui pourrait être portée dans un cadre mixte lecture publique-universités par un groupe-projet constitué notamment par la BPI, la BnF et l'Enssib.

Proposition 4 : Constituer un groupe-projet sur l'observation des usages liés au numérique.

La Commission Bibliothèques numériques envisage de faciliter la constitution en 2012 de ce groupe-projet.

3. La numérisation

Des programmes de numérisation de plus en plus nombreux sont engagés à des échelons divers, national, régional, local ou universitaire, par des bibliothèques ou par des instituts et centres de recherche, pour des finalités multiples, qu'il s'agisse de la valorisation du patrimoine, du soutien à la recherche ou de l'action éducative et culturelle auprès de larges publics, dans le cadre scolaire ou à l'échelle d'un territoire. Certains de ces programmes combinent différentes logiques, territoriales ou thématiques, comprenant parfois en même temps des volets éducatifs et culturels, patrimoniaux ou de recherche, et des coopérations à de multiples niveaux. Les perspectives de mise en valeur concernent aussi bien la médiation auprès d'un public de jeunes ou d'amateurs, qu'une valorisation très clairement orientée vers la recherche. Certains participent à des projets européens ou internationaux, d'autres couvrent des périmètres plus larges que les bibliothèques et concernent également les archives et les musées. Mais, si on constate l'émergence en nombre de projets de numérisation ambitieux pour les bibliothèques publiques ou universitaires, il n'est pas aisé d'identifier les ressources susceptible d'intéresser telle ou telle communauté d'usagers.

Par ailleurs, si les projets se sont portés dans un premier temps sur les œuvres du domaine public, ils abordent maintenant la numérisation des œuvres sous droit et leur circulation dans le respect du droit de propriété intellectuelle. Valoriser les bibliothèques numériques et en développer l'usage c'est aussi en effet régler la question de **l'accès aux contenus sous droit dans les bibliothèques**, qui se pose différemment pour les ouvrages épuisés et indisponibles et pour les ouvrages disponibles. Pour les ouvrages disponibles, un accès sera proposé aux bibliothèques selon un modèle économique à définir. Pour les ouvrages indisponibles, un débat est ouvert sur les conditions de l'offre aux bibliothèques dans un système de gestion collective des droits tel que fixé par la loi adoptée par le Parlement relative à l'exploitation numérique des livres indisponibles du XXe siècle, qui concerne les œuvres orphelines et les œuvres épuisées. En effet, l'exploitation de ces œuvres se fera soit par les éditeurs de la version imprimée eux-mêmes soit à travers une société de projet qui sera chargée de la numérisation des œuvres et pourra disposer de licences reçues de la société de gestion collective. Elle devra élaborer un modèle économique viable en fonctions des investissements réalisés, notamment dans le cadre des Investissements d'avenir. On pourrait ainsi envisager, à partir du moment où ces corpus auront été financés par des avances et non par des subventions, que le secteur de l'enseignement supérieur et de la recherche s'en porte acquéreur.

C'est la BnF qui a été chargée par la loi du 1er mars 2012 de la mise en œuvre de la base signalant les œuvres indisponibles à numériser, une œuvre pouvant avoir un certain nombre de manifestations, certaines épuisées, d'autres non. Il n'est pas envisageable en effet que passe dans tous les cas en gestion collective un ouvrage qui serait encore commercialisé dans une autre édition que celle qui est épuisée. C'est un travail de croisements et de vérifications entre les bases existantes (aussi bien les bases des ayants droit que les catalogues de bibliothèque) qui doit fournir les listes bibliographiques les plus fiables possible sur lesquelles les ayants droit auront un droit de retrait. Celui-ci pourra s'exercer soit de façon définitive au titre du droit moral pour les auteurs, soit de façon ponctuelle pour les éditeurs, à charge pour eux d'exploiter les ouvrages sur un mode contractuel dans un délai de deux ans, faute de quoi ils repasseront en gestion collective. Il faut également prendre en compte l'importance de la dimension de numérisation du projet. Dans la mesure où le corpus est potentiellement de plusieurs centaines de milliers de titres, il s'agira d'un programme de numérisation de masse.

Dans le cadre de l'enseignement supérieur et de la recherche, la BSN, qui est avant tout un programme pour les scientifiques, prendra en compte en premier lieu le caractère scientifique des projets et non leur caractère patrimonial. Les projets financés par la BSN devront répondre à des critères de qualité, à la fois scientifique mais aussi technique, tout en garantissant l'archivage des fichiers de manière optimale. L'accent sera mis également par la BSN sur le développement d'outils informatiques pour la construction des bibliothèques numériques réutilisables dans tous les projets de numérisation par tous les organismes et institutions qui seraient intéressés, sans que soit obligatoirement mise en place une interface de diffusion commune pour l'ensemble de ces institutions. L'articulation entre bibliothèques numériques ainsi que la réutilisation de données devraient être facilitées, et permettre via le web de données des liens automatiques à échelle industrielle entre la production numérisée et les catalogues, à la place de liens effectués manuellement. Le développement d'outils communs permettrait également d'harmoniser les pratiques et de garantir l'interopérabilité des plateformes. La BSN vise également à développer ou renforcer les relations avec le MCC et la BnF, notamment pour l'enrichissement, par les projets de recherche, des contenus déjà diffusés sur Gallica, ainsi que pour l'exploitation de Gallica par les chercheurs. Pour la BSN, il conviendrait également d'étudier les modalités de collaboration avec le MCC pour la numérisation du patrimoine documentaire conservé dans les bibliothèques de l'enseignement supérieur, ainsi que par les laboratoires de recherche, à de fins de préservation de cette mémoire.

Dans l'enseignement supérieur et la recherche, le numérique entraîne une prolifération de projets de caractère extrêmement centrifuge qui invalident toute tentative de centraliser à un niveau ou à un autre, quel que soit le bien-fondé en termes d'économie et de circulation de l'information, comme le montrent les difficultés de NUMES, outil de recensement des corpus numérisés et des projets de numérisation. Au plan scientifique, de multiples projets liés à tel ou tel programme de recherche émergent à un moment donné. Pour ce qui est de la logique patrimoniale, chaque bibliothèque essaie d'agir en fonction de son propre fonds patrimonial ou des possibilités de subventions sur tel ou tel critère. Certaines numérisent des documents qui le sont déjà quand ils sont au cœur de l'identité de leur projet. Ces limites à la centralisation rendent d'autant plus nécessaire de progresser dans l'interopérabilité des outils de façon à permettre à chacun d'échanger quel que soit son prisme. La BSN n'envisage pas de portail commun de l'ensemble des corpus ou des produits numérisés de l'enseignement supérieur et de la recherche. Pour les revues scientifiques en SHS, PERSEE est le portail de référence des collections antérieures à notre siècle ; pour les textes imprimés et soumis au dépôt légal, Gallica s'est affirmé comme le portail de référence pour les chercheurs. Au-delà d'un simple espace de coopération dans Gallica, des complémentarités pourraient être affichées avec la BSN, dans un Gallica Sciences, qui donnerait de la visibilité aux corpus numérisés issus de l'enseignement supérieur et la recherche. Il faut en même temps avancer dans le soutien à des projets de recherche, dans le travail de conviction auprès des chercheurs pour leur faire comprendre que leur travail sur les corpus a vocation à circuler, et dans le soutien aux fédérations qui se mettent en place comme en linguistique et littérature. Pour faciliter l'exploitation des corpus par les utilisateurs, les bibliothèques doivent également mettre en oeuvre des outils de recherche et de découverte aux fonctionnalités plus avancées que celles des catalogues actuels.

Dans un contexte multipolaire, où il n'existe pas de centre unique, il s'agit donc de renforcer la connaissance partagée, la coopération et l'interopérabilité des différentes bibliothèques numériques, d'accroître leur visibilité et d'encourager les usages les plus diversifiés par les différentes communautés, tout en ayant le souci d'une gestion économe et responsable des deniers publics. *De ce fait, les bibliothécaires ont une grande responsabilité*

dans le recensement de ce qui existe, ainsi que dans la sensibilisation à ces ressources et dans la médiation. Par ailleurs, le public est particulièrement demandeur de toutes les ressources, notamment iconographiques, attachées à des ouvrages précieux. Tel ou tel aspect prédomine selon le prisme propre à chaque projet, et il serait utile de rendre visible à tous les proximités et les complémentarités, de faire connaître les périmètres, les thématiques et les critères de sélection. **Les questions de visibilité, de signalement, d'accès d'une façon générale sont primordiales.**

Cette multipolarité à la base rend d'autant plus nécessaire la normalisation technique et la connaissance partagée des programmes de numérisation, de façon à rendre le plus possible compatibles les projets de numérisation de masse dans une logique patrimoniale et les projets de numérisation de niche dans une logique de recherche, à procéder à la constitution de corpus réutilisables par d'autres dans *un va-et-vient entre masse et niche*. Les logiques ne sont pas toutefois pleinement homogènes : il existe des offres, des corpus et des projets qui sont liés à un financement national historique. Chaque projet doit expliciter ses critères et ses contenus.

La réutilisabilité des oeuvres numérisées dans les différents projets nécessite de faciliter l'interopérabilité et le moissonnage dans un sens comme dans l'autre, y compris le *moissonnage à façon*, les objets étant numérisés en fonction des besoins des populations d'utilisateurs. Pour la nécessaire homogénéité technique, la qualité des métadonnées est importante. Dans certains projets, leur création ou leur enrichissement par les scientifiques sont précieux.

Malgré les contraintes de sécurité informatique, Gallica devrait adopter un mode de fonctionnement plus coopératif, notamment pour faciliter la sélection à volonté et la constitution de sets OAI par les bibliothèques partenaires elles-mêmes. Plus globalement, le système est appelé à être plus ouvert pour répondre au désir d'une plus grande autonomie des établissements partenaires dans la constitution de leurs propres bibliothèques numériques.

Le recours à Europeana comme réservoir de métadonnées permettrait de fédérer diverses initiatives et de mieux structurer l'organisation du moissonnage en relation avec les réservoirs européens. Cela demande toutefois de clarifier les questions juridiques de réutilisation des métadonnées dans les différents secteurs culturels, l'autorisation de toutes les utilisations commerciales et non-commerciales pouvant poser problème par exemple pour des documents d'archive où des tiers possèdent des droits. A cet égard, le rapport *“La réutilisation des informations du secteur public culturel”*, remis au ministre de la culture et de la communication par François Braize en février 2012, dresse un état des lieux et présente une analyse du cadre juridique actuel et de ses principales évolutions en cours au niveau européen. Cette question des droits se pose également pour la recherche, dont le point de vue est que lors de la création d'un objet numérique, il faut avoir assuré dès l'abord le cadre légal de la diffusion ultérieure. Il faut progresser vers des solutions qui concilient le droit d'auteur (et la considération de l'auteur) et les besoins de circulation scientifique. A ce sujet, une étude a été menée à l'échelle européenne sur le statut légal des objets qui sont créés dans les activités de recherche (bases de données, etc...). Globalement, cette étude, qui s'inscrit dans les problématiques de réutilisation des données produites par la recherche, réaffirme le principe de la libre circulation des données de recherche dans une volonté d'ouverture.

La question des droits pose aussi celle, traitée de façon variable selon les licences, de la réception des données à des fins de traitement et de consultation avec des moyens propres, question vitale pour la constitution de la plateforme ISTEEX. En effet, certaines licences n'autorisent pas de traitement automatisé sur les données en texte intégral et sur les métadonnées, et ce problème est lié aux formats, protégés et non exploitables. L'ABES est en train de mener une étude juridique concernant la réexploitation des métadonnées.

Proposition 5 :

Définir les conditions d'un équilibre satisfaisant entre propriété intellectuelle et circulation des données et métadonnées numériques pour les bibliothèques, conciliant les logiques de financement et la libre circulation des données publiques.

Pour avancer en ce sens, il convient de créer **un groupe technique de prescriptions**, qui fasse bénéficier la communauté des expériences acquises et préconise des règles d'harmonisation et des normes de qualité ainsi que des procédures d'exploitation légale des métadonnées. Pour être suivi, le cadre de prescriptions techniques doit prendre en compte le fait que les exigences en matière de qualité de numérisation peuvent varier selon les objectifs. Un tel groupe contribuerait également à faciliter la création d'espaces coopératifs, notamment pour améliorer la connaissance, par l'ensemble des acteurs concernés, des projets de numérisation.

Proposition 6 : Créer un groupe technique de prescriptions pour les bibliothèques numériques.

La Commission Bibliothèques numériques envisage de faciliter la mise en place en 2012 de ce groupe.

Le Président de la Commission

Pierre CARBONE

PROPOSITIONS DE LA COMMISSION BIBLIOTHEQUES NUMERIQUES:

Proposition 1 : Mener en 2012 une étude des modèles économiques de tiers-archivage public.

La Commission Bibliothèques numériques envisage de mener cette étude dans le cadre de la poursuite de ses travaux

Proposition 2 : Préfigurer un centre de compétences mutualisé en s'appuyant sur le groupe Pérennisation de l'information numérique (PIN).

La Commission Bibliothèques numériques constituera en 2012 en son sein un groupe de travail réunissant les représentants du MCC, du MESR, de la BnF et du CNRS sur ce sujet et prendra l'attache du SIAF. Il s'agira d'en articuler les travaux avec ceux de la DISIC

Proposition 3 : Faire le bilan des expérimentations d'achat de ressources électroniques accessibles en ligne aux publics inscrits dans les bibliothèques territoriales

La Commission Bibliothèques numériques envisage de travailler à ce bilan en 2012 dans le cadre de la poursuite de ses travaux.

Proposition 4 : Constituer un groupe-projet sur l'observation des usages liés au numérique.

La Commission Bibliothèques numériques envisage de faciliter la constitution en 2012 de ce groupe-projet.

Proposition 5 : Définir les conditions d'un équilibre satisfaisant entre propriété intellectuelle et circulation des données et métadonnées numériques pour les bibliothèques, conciliant les logiques de financement et la libre circulation des données publiques.

Proposition 6 : Créer un groupe technique de prescriptions pour les bibliothèques numériques.

La Commission Bibliothèques numériques envisage de faciliter la mise en place en 2012 de ce groupe.

Annexes

- Annexe 1 : Lettre de mission
- Annexe 2 : Recommandations principales du rapport Schéma numérique des bibliothèques (décembre 2009)
- Annexe 3 : Synthèse des auditions

MINISTÈRE DE L'ENSEIGNEMENT
SUPERIEUR ET DE LA RECHERCHE

MINISTÈRE DE LA CULTURE
ET DE LA COMMUNICATION

8 - NOV. 2010

Monsieur Pierre CARBONE
Conservateur général
Inspection générale des bibliothèques
61-65, rue Dutot
75732 PARIS CEDEX 15

Nos réf. : CC/1489/MMI

Monsieur le Conservateur général,

Le 22 mars dernier, Monsieur Bruno Racine, président de la Bibliothèque nationale de France, a remis le rapport « Schéma numérique des bibliothèques », élaboré par un groupe interministériel associant une vingtaine de représentants du secteur des bibliothèques et du livre : représentants des ministères chargés de l'enseignement supérieur et de la culture, de la Bibliothèque nationale de France, des associations professionnelles de bibliothécaires, etc.

Ce rapport dresse un état des lieux inédit des collections, des services et des usages numériques en bibliothèque. Il constitue ainsi un socle intéressant, tant pour le plan pour le renouveau des bibliothèques universitaires lancé en février 2010 par le ministère de l'Enseignement supérieur et de la Recherche, que pour les 14 propositions pour le développement de la lecture présentées par le ministère de la Culture et de la Communication en mars 2010.

Nos deux ministères ont décidé de continuer à travailler en étroite concertation sur ce sujet. La première recommandation du rapport de Monsieur Bruno Racine consistait d'ailleurs à poursuivre une telle coopération à travers un « groupe de contact permanent sur les bibliothèques numériques ».

Compte tenu de votre implication dans les précédents travaux du rapport « Schéma numérique des bibliothèques » et de votre expérience professionnelle, nous souhaiterions vous confier le soin de présider une commission « Bibliothèques numériques », composée à parité de représentants désignés par nos deux ministères et chargée d'étudier les suites à donner aux 11 recommandations formulées dans le rapport.

.../...

Pour chacune des thématiques suivantes, cette commission pourrait en particulier :

- (ressources électroniques) analyser les expérimentations récentes et le développement de l'offre française de contenus en ligne, définir des cadres d'action adaptés aux spécificités de chaque catégorie de bibliothèque et proposer des modalités de soutien de l'État à ces nouvelles politiques d'acquisition ;
- (numérisation) évaluer les programmes en cours et les modes de collaboration retenus, identifier les nouveaux projets communs ;
- (conservation numérique) préciser les conditions de mise en œuvre d'une politique partagée de conservation numérique entre grands acteurs publics ainsi que l'offre de services aujourd'hui envisageable ; définir les différentes stratégies de conservation des données numériques à mettre en œuvre par les bibliothèques en fonction de leurs objectifs et de leurs moyens ;
- (évaluation) identifier des indicateurs communs entre les différents outils statistiques existants ; proposer les critères d'une évaluation systématique des services numériques et des programmes de numérisation.

Vous déterminerez le plan de travail et les modalités d'action qui vous paraîtront les plus adéquats. La mission de l'information scientifique et technique et du réseau documentaire (service de la coordination stratégique et des territoires) du ministère de l'Enseignement supérieur et de la Recherche et le service du livre et de la lecture (direction générale des médias et des industries culturelles) du ministère de la Culture et de la Communication vous assisteront pour assurer le secrétariat de la commission. Vous voudrez bien nous faire part d'un bilan de vos travaux et de vos recommandations sous la forme d'un court rapport de synthèse annuel qui pourra être présenté devant le Conseil du livre ou toute autre instance interministérielle compétente en la matière.

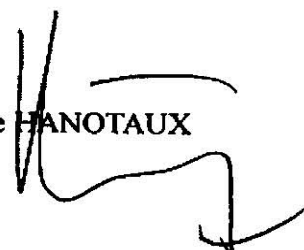
Nous vous prions d'agréer, Monsieur le Conservateur général, l'expression de nos sentiments les meilleurs.

Le Directeur du cabinet du ministre
de l'Enseignement supérieur
et de la Recherche



Olivier PAGEZY

Le Directeur du cabinet du ministre
de la Culture et de la
Communication



Pierre HANOTAUX

Annexe 2

Recommandations principale du rapport Schéma numérique des bibliothèques (décembre 2009)

Recommandation 1 : Pérenniser la concertation autour du numérique en bibliothèques à travers un groupe de contact permanent sur les bibliothèques numériques.

Recommandation 2 : Mettre en œuvre une politique d'acquisition des ressources documentaires électroniques sous forme de licences nationales dans les cas qui s'avèreront les plus appropriés.

[♣ Public-Privé]

Recommandation 3 : Créer un consortium national (GIP, association, fondation ou autre forme) pour coordonner la conservation numérique en bibliothèque, le cas échéant dans d'autres institutions publiques (institutions de recherche, institutions culturelles (archives, musées) ou documentaires, etc.).

Recommandation 4 : Fournir à leurs instances de décision ainsi qu'à leurs tutelles, les critères d'une évaluation systématique des services numériques et des programmes de numérisation mis en œuvre par les bibliothèques.

Recommandation 5 : Développer le recensement des programmes de numérisation par une mise en réseau renforcée des bases d'inventaire existantes (Patrimoine numérique, NUMES, RNBFD).

Recommandation 6 : Constituer les corpus documentaires numériques les plus exhaustifs possible dans des domaines de référence répondant à la demande du public et à des nécessités de préservation du patrimoine, grâce à des programmes coopératifs de numérisation ouverts à toutes les bibliothèques.

Recommandation 7 : Définir et financer un programme concerté et d'envergure de numérisation de la presse nationale et régionale des XIXe et XXe siècles. Lancer une action nationale pour rassembler et valoriser les collections de presse nationale et régionale déjà numérisées.

Recommandation 8 : Aider les bibliothèques à mettre en place des réservoirs de données répondant au protocole OAI-PMH. En faire un critère majeur de la contractualisation en cas d'attribution de subventions publiques pour la numérisation.

Recommandation 9 : Développer l'offre légale de livres numériques, dans le cadre d'un plan de développement du livre numérique en bibliothèque conçu en partenariat avec les éditeurs, et rendre plus visibles ces ressources.

[♣ Public-Privé]

Recommandation 10 : Développer l'offre de contenus numériques pour les handicapés, et offrir un signalement exhaustif de cette offre à travers un outil de signalement national comme le Catalogue collectif de France.

Recommandation 11 : Inciter les organismes de formation à développer une offre de formation initiale et continue répondant aux besoins des bibliothèques de toute taille et mieux faire connaître cette offre.

Lien vers le rapport *Schéma numérique des bibliothèques* :

<http://www.ladocumentationfrancaise.fr/rapports-publics/104000143/>

[♣ Public-Privé] : Indique les actions pouvant s'appuyer sur des partenariats public-privé.

Annexe 3

SYNTHESE DES AUDITIONS

Le texte qui suit n'est pas un compte-rendu *verbatim*, mais une synthèse des auditions, dont le rapporteur assume la responsabilité.

1. L'archivage pérenne des documents numériques (Séance du 30 juin 2011)

Audition de Jean-Charles Hourcade (groupe Pérennité des supports numériques commun à l'Académies des Sciences et à l'Académie des Technologies)

Les documents produits ou acquis sous forme numérique connaissent une croissance exponentielle. La pérennité des contenus numériques est liée d'abord à la pérennité des supports. La durée de vie des supports optiques étant limitée, les bandes magnétiques restent encore pour l'électronique le meilleur support en termes de longévité et de rapport coût/qualité. Cette durée de vie limitée (10-15 ans) contraint à des migrations périodiques de supports. D'autres technologies innovantes ont actuellement un coût de production qui ne permet pas leur commercialisation. Le passage au numérique étant définitivement installé, il convient de privilégier pour les bibliothèques numériques une approche multi-technologies et multi-supports, compte tenu des coûts élevés d'investissement et de fonctionnement (moyens humains compris) pour le stockage. Les coûts de stockage, qui sont actuellement en moyenne de plusieurs milliers d'euros par téraoctet (To) et par an, devraient toutefois baisser à l'horizon 2015. La mutualisation de l'archivage comme sa gestion sur la longue durée sont indispensables.

Audition de Françoise Banat-Berger (Direction des Archives de France)

Pour la pérennité des contenus, les procédures d'archivage ne se limitent pas à un simple stockage de données et doivent s'appuyer sur une documentation des processus métier. Celle-ci doit prendre en compte la gestion des risques, garantir l'application des normes et standards dans le respect de l'interopérabilité, et assurer la gestion des métadonnées. Notamment, les métadonnées doivent être gérées dès le début du processus.

Il convient de diffuser les normes et standards existants et former à leur emploi : la norme OAIS ISO 14721:2003, la DTD EAD pour la description des documents d'archives, la norme NF Z 42-013 sur l'archivage numérique (adoptée comme norme ISO 14641-1), ainsi que le format SEDA, standard d'échanges de données entre systèmes d'information pour l'archivage élaboré par l'actuel Service interministériel des archives de France (SIAF). La formation au format SEDA est nécessaire pour accompagner le passage vers des processus automatisés et contrôlés par des machines.

Ces spécifications viennent à l'appui des recommandations *Politique et pratiques d'archivage (P2A)*, arrêtées en 2006 par les Archives de France avec l'ancienne Direction centrale de la sécurité des systèmes d'information (DCSSI) et qui définissent notamment les responsabilités et les rôles des différents acteurs (autorités d'archivage, opérateurs, services de contrôle).

L'archivage électronique peut être externalisé pour les archives courantes et intermédiaires, le prestataire de tiers-archivage devant être agréé par le MCC sur la base d'une grille élaborée par le SIAF. Mais cette externalisation s'effectue dans le cadre de marchés dont la durée limitée est incompatible avec le temps beaucoup plus long de la conservation.

A la suite du rapport Quenet sur les archives de l'Etat, des évolutions s'imposent dans le sens notamment d'une plus grande mutualisation (au niveau de la gouvernance) et de la prise en charge du nouveau chantier que constitue l'archivage numérique définitif. Plus globalement Françoise Banat-Berger plaide pour l'existence d'un centre de compétences mutualisé pour l'ensemble des acteurs s'intéressant à la préservation de l'information numérique, qui prenne en compte les questions de veille sur les formats, les supports, et les formats de métadonnées et de formation, en partenariat avec les autres opérateurs institutionnels concernés par cette question (BnF, CINES, CNES).

Audition d'Arnaud Beaufort, Thomas Ledoux et Laurent Duploux (Bibliothèque nationale de France)

La BnF conserve des documents numériques de nature, d'origine et de statut juridique divers au sein de sa plateforme SPAR (Système de préservation et d'archivage réparti) hébergée sur deux sites, à Tolbiac et à Bussy-Saint-Georges. L'archivage est actuellement organisé en huit filières : Numérisation de conservation, Audiovisuel et multimédia, Numérisation de consultation, Dépôt légal automatique, Dépôt légal négocié, Production administrative et technique, Dons et acquisitions, Tiers-archivage. La volumétrie fin 2011 est de 910 To, la capacité de stockage devant atteindre 4 650 To en 2013. La BnF formalisera son offre de tiers-archivage payant lors d'un conseil d'administration en juin 2012.

La coopération avec d'autres partenaires constitue pour la BnF un enjeu stratégique majeur sur le plan national. De fait, en considérant les préconisations du groupe Racine (SNB) et la constitution de la BSN (Bibliothèque scientifique numérique), Arnaud Beaufort estime que la mise en place d'un groupement indépendant fédérant les acteurs nationaux concernés pourrait permettre de mener des études et des projets, de coordonner les besoins et de piloter la mutualisation de certains services.

Audition d'Olivier Rouchon (Centre informatique national de l'enseignement supérieur)

Le CINES, dédié au calcul intensif, est chargé depuis 2004 d'une mission nationale d'archivage du patrimoine scientifique pour les secteurs de l'enseignement supérieur et de la recherche, la mutualisation des équipements permettant d'optimiser les coûts. Il s'est doté d'une infrastructure numérique d'archivage pérenne, la plateforme PAC en production depuis 2007, qui respecte une procédure d'assurance qualité. Cette plateforme conserve des données aussi bien scientifiques, pédagogiques qu'administratives, notamment les thèses, les revues numérisées de PERSEE, les dépôts en archive ouverte dans HAL, des documents sonores versés par le TGE Adonis ainsi que des ouvrages numérisés par des bibliothèques universitaires. Le CINES s'est engagé dans un projet d'archivage intermédiaire des données primaires scientifiques (le projet isaac), et est également en pourparlers avec l'Institut national de la santé et de la recherche médicale et la Cour des Comptes. Au 1^{er} juin 2011, sur une capacité de 40 To, la volumétrie archivée s'élève à 12 To (hors copie multiple), et elle devrait doubler en 2012. Son modèle économique de tiers-archivage comprend trois catégories : tarifs internes (pour les établissements dépendant du MESR) ; tarifs partenaires (autres établissements d'enseignement supérieur et de recherche hors MESR) ; tarifs externes (les autres établissements hors secteur enseignement supérieur et recherche). Le CINES n'est pas sur ce point en concurrence avec les autres acteurs publics français, encore moins avec les acteurs privés qui s'adressent à une clientèle d'entreprises.

Dans les faits, les actions du CINES en matière d'archivage pérenne du numérique s'inscrivent déjà à la croisée des partenariats, aussi bien au sein de l'ESR (avec le TGE Adonis par exemple), qu'au-delà (avec le SIAF ou la BnF), ainsi qu'au niveau européen (projet EUDAT d'infrastructure européenne de conservation/diffusion de données scientifiques). Pour autant, Olivier Rouchon plaide pour que ces partenariats s'inscrivent dans un cadre formel qui impliquerait les acteurs dans une coopération stratégique au niveau national. Les enjeux de cette coopération formelle sont multiples : coordination des besoins, activités normatives, veille technologie, guides de bonnes pratiques, financement d'études et de projets, mutualisation de services. Par ailleurs elle supposerait une prise en charge financière par l'Etat notamment pour décharger les producteurs de données du poids des coûts d'archivage.

Audition de Christine Berthaud et Daniel Charnay (Centre de communication scientifique directe du CNRS)

Le Centre de communication scientifique directe (CCSD), unité propre de service du CNRS au service de tous les établissements publics scientifiques et techniques, gère essentiellement des archives ouvertes de la recherche au travers d'HyperArticles en ligne (HAL). Le CCSD conserve des documents en différents formats (Word, LaTeX...), dont des formats d'échange comme le PDF. La sécurisation des données est assurée par l'hébergement d'HAL à l'IN2P3 avec une sauvegarde au CINES. Pour l'archivage à long terme, les documents confiés autres que PDF sont systématiquement transcrits dans ce format. La volumétrie traitée s'élève à environ 1 To pour 150 000 documents. Le CCSD traite aussi bien la prépublication que la post-publication.

Audition de Richard Walter (TGE Adonis)

Le Très grand équipement (TGE) Adonis, unité propre de service du CNRS, est une infrastructure numérique en sciences humaines et sociales (SHS). Adonis anime, dans une logique de standardisation des modes de production et de diffusion, un réseau de centres de ressources produisant des objets numériques (texte, image, son, 3D, etc.). Il assure l'accessibilité et la préservation des données numériques à travers une grille de services basée à l'IN2P3, qui offre aux équipes concernées une structure informatique pérenne. Le TGE Adonis propose des services mutualisables, notamment l'hébergement de sites Web, le stockage de données et l'archivage à long terme des données de la recherche. Il gère également la plateforme Isidore de recherche et de diffusion en SHS. Il coordonne ces actions grâce à des partenariats avec l'IN2P3, le CINES et le CCSD. Pour l'archivage à long terme, un projet pilote relatif aux données orales devrait être suivi d'autres projets reposant sur des formats ouverts et documentés.

Audition de François Berquet et Damien Leblanc (Médiathèque du Grand Troyes)

La Médiathèque de Troyes met en œuvre depuis 1998 une politique de numérisation et de valorisation de son patrimoine, et a constitué un service *ad hoc*, qui travaille en lien avec le service informatique. Les fonds numérisés (manuscripts médiévaux dans le cadre d'un programme réalisé par l'IRHT, livrets et almanachs de colportage, incunables, ouvrages anciens) représentent aujourd'hui 1,2 million de pages et 14 000 images accessibles en ligne. La conservation a été réalisée dans un premier temps sur divers supports de stockage. Par la suite, c'est à une société de services informatiques, gestionnaire du portail, qu'a été confiée la gestion des données numériques : des images en haute définition auxquelles sont attachées les métadonnées. Actuellement, 13 To sont utilisés sur une capacité totale de 18, la montée en charge annuelle représentant 1 à 2 To. La question de la capacité de stockage à court terme se pose aujourd'hui d'une manière cruciale, dans la mesure où les prévisions initiales couvraient 10 années de fonctionnement. Dans le cadre des projets menés en partenariat avec la BnF (Bibliothèque numérique de Clairvaux, catalogue collectif des pièces de colportage, programme de numérisation d'imprimés humanistes), la participation à SPAR est envisagée. Plus globalement, la Médiathèque de Troyes ne verrait que des avantages à participer à une plateforme nationale du patrimoine numérisé des bibliothèques.

2. L'acquisition de ressources électroniques (Séance du 12 octobre 2011)

Audition de Serge Bauin (Directeur de l'information scientifique et technique au CNRS)

La coordination des activités d'information scientifique et technique (IST) au sein du secteur de l'enseignement supérieur et de la recherche se renforce avec la mise en place de la Très grande infrastructure de recherche (TGIR) *Bibliothèque scientifique numérique (BSN)*. Le Ministère de la Culture et de la Communication (MCC) siège au Comité de pilotage de la BSN ; il est membre du groupe BSN1 sur l'acquisition des ressources électroniques et la BnF est associée aux travaux de la BSN relatifs aux acquisitions électroniques, à l'archivage pérenne et à la numérisation. La BSN permet un pilotage scientifique des acquisitions, grâce à un dialogue entre les chercheurs et les professionnels de l'IST. La négociation des ressources électroniques doit aussi prendre en compte le programme AIDA de professionnalisation des achats de l'État. Les acquisitions électroniques doivent en outre prendre en compte les unités mixtes de recherche, qui associent des opérateurs desservant la même communauté, ce qui implique entre eux une répartition des coûts et une politique coordonnée d'accès.

Un élément majeur de la BSN est l'acquisition de ressources électroniques en *licence nationale*, qui permet des économies d'échelle. Les projets en cours concernent des archives rétrospectives de périodiques scientifiques et des corpus. L'Agence bibliographique de l'enseignement supérieur (ABES), mandatée par le Ministère de l'enseignement supérieur et de la recherche (MESR) selon les priorités établies par la BSN, a passé de premiers contrats dont des bibliothèques publiques pourront bénéficier, parmi lesquelles la BnF et la Bibliothèque publique d'information (BPI). Il faudra également étendre, selon un modèle économique à définir, les licences nationales aux ressources courantes, le courant d'aujourd'hui constituant l'archive de demain.

Audition de Jacques Fontanille (Conférence des présidents d'université)

Les licences nationales sont au cœur du projet ISTEEX, projet retenu à hauteur de 60 M€ sur trois ans dans le cadre des initiatives d'excellence (IDEX). ISTEEX est un programme d'achat de collections rétrospectives de périodiques et autres ressources numériques, pour constituer une bibliothèque numérique aux meilleurs standards internationaux, instrument accessible à tous les chercheurs pour leurs travaux. Le projet inscrit de façon native une double dimension: acquérir un corpus inégalé de ressources sous forme de licences nationales, et les rendre accessibles et exploitables au sein d'une plateforme nationale. Ce programme, transverse à toutes les IDEX, sera un instrument au service de la recherche française, et permettra des gains en compétitivité. Du fait du changement d'échelle qu'il implique, ISTEEX offrira à l'ensemble des chercheurs l'accès à un corpus inédit de ressources, intéressant toutes les communautés scientifiques. Accessibles à distance par tous les membres des établissements de l'enseignement supérieur et de la recherche, ces ressources seront en outre valorisées par des outils de recherche et des services à valeur ajoutée permettant de les exploiter pleinement (recherches portant sur l'ensemble du corpus, exploration automatisée de textes, enrichissements ...). Un des objectifs de la plateforme est également de développer des outils interopérables, pour permettre aux établissements une appropriation de l'outil et une intégration des ressources dans leurs environnements locaux, facilitant ainsi la navigation des chercheurs parmi l'ensemble des ressources qui leur sont offertes (productions locales, abonnements courants, archives) et de mutualiser les innovations technologiques. ISTEEX est porté par quatre partenaires : l'université de Lorraine, le CNRS, l'ABES et Couperin, et

s'inscrit dans le cadre de la BSN. La convention entre l'Agence nationale de la recherche, l'Etat et le porteur du projet devrait être signée prochainement.

Les statistiques d'usages ont une grande importance dans un tel projet, les données sur les catégories d'utilisateurs des ressources pouvant éclairer les politiques d'achat. Il faudra veiller à l'articulation entre le niveau national et le niveau local pour l'analyse des données d'usages.

Audition d'Hazel Woodward (Cranfield University, Joint Information Systems Committee – JISC, Royaume Uni)

A côté de ces vastes corpus, les achats de livres électroniques sont encore au stade expérimental, y compris pour les manuels. Les offres sont hétérogènes : achat pérenne de titres uniques, collections en achat pérenne ou en abonnement, abonnement à une base de contenus agréant les catalogues de plusieurs éditeurs, collections de contenus historiques numérisés, collections de référence. Mais l'intérêt majeur des bibliothèques pour les manuels numériques se heurte à un modèle économique fondé sur l'achat individuel par l'étudiant. Face à ces difficultés, au Royaume-Uni, le JISC a créé un observatoire national du livre numérique, qui a testé pendant deux ans dans toutes les universités britanniques l'achat de licences de manuels numériques très demandés, et a évalué les usages ainsi que l'impact de cette offre sur les différents acteurs. Si 61% des étudiants ont consulté (en majorité sur écran) un livre numérique, seuls 47% se le sont procuré à la bibliothèque. Les *e-books* ont maintenant un rôle central dans l'apprentissage, malgré des pratiques de lecture différentes du papier (un temps de lecture plus court, des usagers moins concentrés, une lecture moins extensive...). En même temps qu'ils accèdent à distance à la bibliothèque virtuelle, les étudiants continuent à fréquenter la bibliothèque physique, notamment pour emprunter. Enfin, le signalement des livres numériques (via le catalogue et le site de la bibliothèque) est important pour en favoriser l'usage. Ce test a permis de constater que le modèle dominant, parallèlement à l'offre en bibliothèque, demeurera celui de l'achat individuel par l'étudiant (qui va se déplacer de l'imprimé au numérique et privilégier des contenus lisibles sur des appareils mobiles). La bibliothèque peut fournir des *services associés*, combinant les ressources gratuites de formation en ligne et les manuels numériques payants : en effet, au fur et à mesure que le corpus de textes numériques ou numérisés augmente, les liens entre les documents augmenteront et légitimeront leur usage, avec la possibilité pour les lecteurs d'indexer, d'annoter, d'agréger des contenus provenant de livres numériques différents.

Audition de Vincent Piccolo (Editions la Martinière, Commission numérique du Syndicat national de l'édition)

L'offre en livres numériques des éditeurs français à destination des bibliothèques est proposée à travers plusieurs plateformes. Le livre numérique ne représente encore que 1% de leur chiffre d'affaires et le rythme d'évolution est peu prévisible. Pour plusieurs raisons (loi sur le prix du livre numérique, atomisation de la diffusion de livres en France, etc.), le développement de ce marché pourrait y être plus lent que dans les pays anglo-saxons. Les éditeurs ne souhaitent pas adopter une démarche verticale qui les conduirait à se mettre directement au contact du lecteur, mais ont au contraire le souci de maintenir la même offre à travers un réseau d'intermédiaires (la grande librairie, la librairie indépendante et depuis peu, les « *pure players* »). Les éditeurs veulent s'adresser au marché des bibliothèques non pas directement mais via les libraires, pour qui le marché des institutions demeure important. Mais l'hétérogénéité des bibliothèques empêche l'adoption d'un modèle unique. Des rencontres ont d'ores et déjà été organisées avec COUPERIN et CAREL. Pour construire cette offre, il faut développer une brique technique et un dossier a été déposé en ce sens au Centre national du livre (CNL). Le modèle retenu n'est pas celui de la marque blanche (procédé par

lequel une entreprise propose à ses clients un service assuré par un tiers sans que celui-ci apparaisse clairement comme en étant le fournisseur), mais celui du *web service*, afin de mettre en place un système permettant aux libraires de répondre aux appels d'offres des bibliothèques. Une expérimentation est déjà en cours avec la BanQ (Bibliothèque et archives nationales du Québec). En ce qui concerne les bibliothèques, les éditeurs souhaitent proposer un modèle d'affaires proche de celui du livre papier. Ce modèle, qui serait adapté progressivement dans un processus de codéveloppement, pourrait être le suivant dans un premier temps :

- l'achat pérenne et des consultations en *streaming* illimitées dans l'emprise de la bibliothèque, avec une remise de 20% par rapport au prix du livre papier,
- l'accès distant par téléchargement, avec des mesures techniques de protection (DRM) basées sur la chronodégradabilité aussi souples et standardisées que possibles.

Si dans le secteur académique, l'offre française est déjà plus souple que l'offre étrangère, le marché de littérature générale diffère de celui des bases de données scientifiques, et le modèle de la sélection à l'unité pourrait être préféré à celui du bouquet. Le développement de l'offre, avec l'accord des ayants droit, exige le respect des DRM. Une offre illimitée accessible à distance en *streaming* représenterait un manque à gagner trop important pour les éditeurs, qui sont en revanche conscients de la nécessité de mener des expérimentations avec les bibliothèques.

Audition de Michel Fauchié (Association pour le développement des documents numériques en bibliothèque)

Les usages des bibliothèques publiques en matière de ressources numériques montrent que les lecteurs utilisent largement les accès publics à Internet dans les bibliothèques. Mais l'offre de contenus numériques y est encore inadaptée – notamment en raison des nombreuses contraintes techniques que les fournisseurs imposent – et sous-utilisée. Pour les bibliothécaires, les grands éditeurs hésitent à leur proposer une offre numérique large, qui leur apparaît faire concurrence à l'achat individuel. A contrario, les petits éditeurs ou de nouveaux distributeurs cherchent à se positionner sur le marché, mais les conditions d'utilisation ou les contenus proposés sont limités. Les bibliothèques souhaitent des offres associant des documents susceptibles d'intéresser un large public, un accès ergonomique et harmonisé aux ressources, interfaçables dans un catalogue unique, avec des coûts d'abonnement maîtrisables. Elles souhaitent négocier avec les éditeurs des modalités d'accès, par exemple le *streaming*, bien accueilli pour la musique et la vidéo, et qui pourrait s'étendre au livre. Il faudrait également expérimenter de nouveaux modèles de livres numériques, autres que le livre homothétique. Au-delà du prix, les principaux obstacles pour les bibliothèques et leurs utilisateurs sont les difficultés d'accès, d'authentification et de signalement. On note cependant un certain nombre d'expériences en cours dans des bibliothèques territoriales, proposant des ressources à l'ensemble de la population (presse, musique en ligne, VOD) dont il serait utile de tirer le bilan. Mais les obstacles en bibliothèque publique sont encore importants : l'accompagnement des personnels et des usagers est insuffisant, les infrastructures techniques posent souvent problème, et les ressources ne sont pas suffisamment visibles. Surtout, les usagers attendent un accès beaucoup plus simple aux ressources. Il en résulte que, si un nombre croissant de bibliothèques territoriales s'engagent dans l'achat de ressources numériques, la majorité reste attentiste devant l'instabilité des modèles.

Avec la baisse de l'emprunt de certains documents, les bibliothèques sont confrontées aux changements de pratiques et à l'hybridation des usages entre documents physiques et numériques. Mais au plan des moyens, le numérique représente un coût important en matériel et en abonnements, et demande également des compétences, notamment techniques, mais

aussi en matière de communication et de médiation. Il s'agit aussi d'expérimenter et d'analyser les usages à partir d'indicateurs quantitatifs et qualitatifs. Si certaines bibliothèques territoriales ont les moyens d'acheter certains abonnements disponibles sur le marché, seules les plus volontaristes vont jusqu'à construire des projets propres, liés à des contenus locaux (patrimoniaux ou contemporains) ou à des actions de médiation numérique.

Audition d'Annie Dourlent (CAREL, Bibliothèque publique d'information)

L'organisation d'un réseau de veille, de sélection et de négociation de ressources électroniques pour les bibliothèques territoriales a été préfigurée par CAREL. Créé il y a huit ans au sein de la BPI, CAREL a joué un rôle très positif pour faire connaître les ressources, proposer des tarifs de référence et constituer un réseau. Depuis 2005, 243 établissements ont été recensés par CAREL. En 2010, 115 bibliothèques de lecture publique ont répondu à son enquête sur les abonnements aux ressources numériques payantes, dont 95 ont déclaré s'être abonnées par le biais de CAREL. Le catalogue proposé par CAREL contient une soixantaine de ressources, sélectionnées sur la base de l'expertise des acquéreurs de la BPI et en lien avec les établissements de lecture publique. Collectant des informations au niveau national, il offre la possibilité aux fournisseurs de s'adresser à un interlocuteur bien identifié. Mais il n'y a pas de procédure d'adhésion ni d'enquête sur les intentions d'achat, et de ce fait les éditeurs ne s'engagent pas sur un tarif. De plus, les moyens actuels de CAREL ne lui permettent pas d'agir à plus grande échelle. Par ailleurs, les usages demeurent relativement faibles, sauf pour les ressources de presse en ligne et d'autoformation. Pour CAREL, un autre obstacle tient à la méconnaissance des bibliothèques par les éditeurs et les agrégateurs de ressources numériques, qui pratiquent souvent des grilles tarifaires peu réalistes. Néanmoins, de petits fournisseurs sont désormais dans une logique de co-construction de modèles avec CAREL. Les bibliothèques souhaitent expérimenter de nouvelles ressources, notamment multimédias, elles attendent aussi un outil collaboratif en ligne permettant un meilleur partage de l'information, des modèles économiques plus adaptés. CAREL pourrait laisser la place à une association de collectivités territoriales, où la BPI – et à travers elle, l'Etat – continuerait à jouer un rôle de premier plan. Pour aider les bibliothèques dans leurs réflexions, échanger des informations, des bonnes pratiques, des savoir-faire, repérer et conseiller les meilleures ressources électroniques, orienter, aider à négocier de façon informée. Ce véritable outil professionnel de mutualisation s'appuyant sur un réseau d'experts et de médiateurs permettrait de négocier les ressources numériques dans de meilleures conditions.

3. La numérisation (Séance du 1^{er} décembre 2011)

Audition d'Arnaud Beaufort (Bibliothèque nationale de France)

A la BnF, Gallica a évolué d'une démarche très sélective (5 000 à 7 000 titres numérisés par an avant 2005) à un chantier de numérisation de masse des imprimés à partir de 2007 (avec un marché dit « des 100 000 » volumes par an), couplé à la numérisation des collections spécialisées (manuscrits, estampes, photographies, monnaies et médailles...). Ce programme, moissonné par Europeana, concerne uniquement les collections françaises, à l'exclusion des documents trop fragiles. Il porte sur les fonds suivants : la totalité des collections de livres français (environ 8 millions dont 3 millions dans le domaine public) ; la numérisation en priorité de la presse (en particulier pour la période 1870-1940 dont le papier se dégrade et qui est estimée à 100 millions de pages) ; les documents spécialisés selon une exhaustivité raisonnée. Au rythme actuel, la numérisation des seuls ouvrages hors droit demanderait 30 ans, ce qui justifie la recherche de partenaires privés. Quant aux œuvres sous droit (500 000

ouvrages), des actions sont envisagées : une loi a été votée le 22 février 2012 et promulguée le 1^{er} mars 2012 pour permettre la numérisation des œuvres qui ne sont plus disponibles dans le commerce. La mise en œuvre du dépôt légal numérique à venir éviterait que le stock à numériser rétrospectivement n'augmente de 70 000 à 80 000 ouvrages par an. Les contenus déjà numérisés par la BnF concernaient jusqu'alors en priorité l'histoire de France, la littérature et les langues, ainsi que des corpus numérisés intégralement dans le cadre de coopérations.

A été mis également en place en 2011 avec l'Association des Amis de la BnF le programme « *Adopter un livre* », qui fait appel à des donateurs pour des titres dont la numérisation n'est pas prévue dans l'immédiat, une mention laissée au choix du donateur étant associée pendant dix ans à l'ouvrage.

Tout internaute peut aussi se réapproprier des données et partager ses coups de cœur Gallica par export sur les blogs, les murs Facebook, etc.

L'augmentation du nombre de documents disponibles dans Gallica (1,7 million de documents à ce jour) a des effets directs sur la progression des usages. En 2011, on comptait 1,5 millions de visiteurs, plus de 9 millions de visites, et la fréquentation atteint désormais 32 000 visites par jour (avec des pics jusqu'à 50 000). Mais la part numérisée de la collection physique totale en français est encore très faible, ce qui justifie de faire porter l'effort public sur la numérisation, et de se tourner en complément vers les entreprises prêtes à accompagner la BnF.

La BnF a constitué trois filières de numérisation :

- *les ateliers internes* sur les sites Richelieu, Bussy-Saint-Georges et Sablé qui réalisent deux millions de clichés par an pour des documents rares, précieux, fragiles, abîmés ou nécessitant une restauration ainsi que pour les grands formats. Les tables des matières, la mise au format ePub et la reconnaissance optique de caractère (OCR) font l'objet de marchés complémentaires afin de mettre en ligne également du mode texte ;
- *de grands marchés avec des prestataires externes*, organisés selon une logique industrielle garantissant le suivi constant de flux d'échange importants ainsi qu'une traçabilité à tout instant des documents, et permettant la numérisation d'une dizaine de millions de pages par an pour les livres et imprimés, la presse et les collections spécialisées ;
- *un appel à partenariats dans le cadre des investissements d'avenir* publié en 2011 et qui portait sur douze corpus : le livre ancien de 1470 à 1700 ; des imprimés français du domaine public ; la presse française de 1780 à 1940 ; les disques 78 tours et microsillons ; les manuscrits médiévaux ou modernes ; les collections de portraits ; la généalogie et l'histoire des familles ; les territoires : la cartographie et les représentations de la France et des pays étrangers ; l'histoire locale ; les partitions musicales ; la photographie ; le cinéma français des origines à la deuxième guerre mondiale. Ce recours à la numérisation par des partenaires privés permettrait de multiplier par six le volume annuel traité. Trois de ces corpus ont dans un premier temps retenu l'attention (imprimés, presse, audiovisuel).

Gallica s'est par ailleurs ouvert à la présentation d'œuvres sous droit dans les emprises de la BnF avec un « *Gallica intra muros* ». Le but est que les lecteurs aient accès à l'ensemble de la bibliothèque numérique de la BnF, sous droit comme libre de droit. Gallica intra muros, qui est d'abord consultable sur les terminaux de la BnF, pourrait être un jour accessible en lecture sur le PC du chercheur lui-même, ce qui nécessite au moins des adaptations techniques pour la protection des fichiers.

Les moyens financiers mobilisés pour les collections imprimées ainsi que pour les collections spécialisées de la Réserve et de l'Arsenal sont à 80% des financements du CNL (12 M€ dont 11,2 M€ pour les livres et imprimés). L'assiette des recettes parafiscales du CNL a été élargie pour financer la participation de la France au développement de la bibliothèque numérique européenne. Pour les autres collections spécialisées, il s'agit de fonds propres abondés pour

certaines programmes, par des crédits européens (1,15 M€ dont 1 M€ en fonds propres BnF). A ces coûts, il faut ajouter ceux qui concernent les marchés d'OCR (1,4 M€ financés entre 50 et 90% par le CNL), ainsi que celui de la presse (0,2 M€ sur fonds propres). Sur la période 2005-2011, la numérisation elle-même ne couvrirait que 37% du coût global, la BnF supportant d'autres charges : sélection documentaire (7%) ; production des métadonnées et contrôle qualité (5%) ; stockage des fichiers (43%) ; programmation et interface (8%).

Le processus de numérisation impacte l'ensemble des métiers de la BnF depuis le prélèvement et l'expédition des documents jusqu'au contrôle qualité et la gestion des fichiers, et y mobilise globalement 250 équivalents temps plein (ETP) - soit 10% du total, ainsi qu'un nombre important de salariés chez les prestataires (le prestataire principal emploie plus de 500 personnes, surtout dans des pays à faible coût de main d'œuvre). En vue des opérations en partenariat, la BnF va de plus créer une filiale qui financera les moyens humains supplémentaires.

Les coûts unitaires au volume peuvent varier dans un rapport de 1 à 10 selon le niveau de qualité exprimé en fonction du taux de reconnaissance de l'OCR souhaité. La numérisation est depuis 2011 réalisée en 400 dpi minimum au lieu de 300 dpi précédemment, en niveaux de gris, ou en couleurs dès qu'au moins un élément est en couleurs. Ce niveau de qualité permettrait de rentabiliser l'effort public de numérisation en développant une offre d'impression à la demande.

La BnF développe également la coopération avec les autres bibliothèques françaises, selon trois modalités :

- le versement dans Gallica de métadonnées de bibliothèques numériques (39 à ce jour), ce qui assure une meilleure visibilité aux 100 000 documents concernés ;
- la numérisation de documents de bibliothèques partenaires choisis en concertation, à raison d'un tiers des volumes traités dans les marchés de la BnF financés par le CNL. Ces bibliothèques reçoivent une copie numérique identique à celle figurant dans Gallica ;
- des téléchargements de fichiers numériques dans Gallica, qui assurent automatiquement leur signalement dans Europeana.

Pour piloter cette coopération, la BnF vient d'ouvrir l'extranet *espacecooperation.bnf.fr*, où les partenaires pourront gérer leur présence dans le portail national tout en ayant la vision du stockage pérenne de leurs fichiers. Cet outil permettra aux bibliothèques de mieux gérer leurs bibliothèques numériques dans Gallica, d'élaborer des plans de travail identifiant pour chaque institution les fonds qui vont être numérisés et intégrés à Gallica dans un esprit de complémentarité qui accélérera la marche vers l'exhaustivité. Par ailleurs, les bibliothèques partenaires y disposeront d'espaces où elles pourront retrouver au moyen de co-marquage ou de marque blanche leur ergonomie, leur logo, et donc une identité propre au sein du portail national.

Audition de Jean-François Sibers (Conseiller Livre et Lecture, DRAC Aquitaine)

A l'échelle régionale, la Banque numérique du savoir en Aquitaine (BnsA) est une des expériences du protocole de décentralisation culturelle. Ce programme de numérisation strictement patrimonial, financé depuis 2000 par deux contrats Etat-Région successifs et soutenu par le Fonds européen de développement régional (FEDER), concerne à la fois les bibliothèques, les archives et les musées. Il associe des collectivités (commune, département, région) et des services déconcentrés de l'Etat (Direction régionale des affaires culturelles – DRAC-, service territorial de l'architecture et du patrimoine) pour créer un service public numérique gratuit à vocation éducative et culturelle en direction des jeunes publics. Chaque collectivité y développe un projet propre à caractère scientifique, culturel et politique (les voyages aux Pyrénées, la Préhistoire dans la vallée de la Vézère, etc...), la diversité des participants et des métiers étant prise en compte dans l'élaboration des outils techniques,

financiers et juridiques. Les partenaires de la BnsA ont des structures administratives différentes (établissement public de coopération culturelle - EPCC, agglomération, “archives de pays” créées par transformation d’une annexe d’archives départementales, etc...). La BnsA n’est pas une nouvelle entité administrative, elle repose sur une charte d’adhésion à caractère politique et financier, qui a permis de mettre à disposition des moyens. Un schéma technique compatible avec les différentes logiques métier a été élaboré, ainsi que des instruments de recherche propres, des fonds significatifs ont été sélectionnés, et la numérisation lancée selon un protocole et une chaîne de traitement. Les compétences des équipes ont été renforcées, au sein des collectivités concernées, notamment en informatique documentaire. Un système d’ingénierie éducative fonctionne depuis 3 ans par convention entre la DRAC, l’Académie et les collectivités productrices de ressources, selon un processus éducatif proposé par le Rectorat aux établissements, et donnant lieu à une formation spécifique.

Le protocole de fonctionnement validé par le MCC vise à ne pas doubler la numérisation par rapport à des réalisations externes ou internes à la BnsA. Les bases nationales du MCC l’alimentent par export, avec la possibilité de les enrichir de ce qui, existant physiquement dans les fonds locaux, n’est ni traité ni visé par elles. Une partie, réservée à l’administration du MCC ou au travail des chercheurs sur le patrimoine, reste consultable dans des espaces restreints en intranet. Les ressources étant dorénavant moissonnables par la BnF, la BnsA va créer des services à valeur ajoutée, notamment une interface cartographique.

Ce programme déborde le cadre régional, avec des projets de numérisation des manuscrits médiévaux en partenariat avec l’Institut de recherche et d’histoire des textes (IRHT) et la BnF et en lien avec Europeana, ou de fonds thématiques avec des musées et centres de recherche. Plusieurs chantiers sont en cours : l’interopérabilité avec Europeana et avec la BnF et pour le catalogue avec le MCC. Les entrepôts étant partiellement moissonnables au niveau local, la BnsA joue le rôle d’un agrégateur régional en mesure de travailler avec des agrégateurs nationaux et internationaux. Dans ce cadre, est à l’étude un protocole coordonné et maîtrisé de moissonnage à façon centré sur des thématiques locales sur la base d’une coopération scientifique.

La BnsA aborde maintenant une nouvelle étape, où il serait pertinent que le programme se pérennise dans une structure administrative (groupement d’intérêt public, EPCC, ...). Elle devrait travailler avec l’Education nationale à des plateformes patrimoniales pour l’enseignement et effectuer un travail systématique de mise à disposition des ressources dans les centres documentaires. Une des questions majeures est la sécurisation des ressources numériques par redondance d’une collectivité à une autre. La question se pose également d’une plus grande ouverture aux partenariats selon un choix à opérer : schéma non marchand et de service public ? partenariat avec des entreprises privées ? schéma mixte ?

Audition de Françoise Blot (Bibliothèque municipale d’Orléans)

A l’échelon local, la Bibliothèque municipale à vocation régionale (BMVR) d’Orléans est la première à avoir obtenu du MCC en 2010 le label de bibliothèque numérique de référence. Ce projet, initié en 2004, s’inscrit maintenant dans le projet d’établissement 2012-2015 de la Ville. La bibliothèque numérique met en ligne sur le site de la BMVR les collections les plus précieuses et fragiles, mais aussi des documents fréquemment consultés en salle, intéressant un public plus large, sur des supports diversifiés et couvrant plusieurs périodes. Pour la presse et les imprimés, la numérisation est effectuée par des prestataires dans le cadre de marchés. Au plan technique, la résolution est à 300 ou 400 dpi avec conversion en mode texte et OCR pour la recherche en plein texte, avec des images en niveaux de gris ou noir et blanc, des fichiers en format standard. Le projet mobilise une équipe de 12 ETP dont 5 cadres A. La BMVR coopère avec la direction des systèmes d’information pour l’archivage des fichiers, la sauvegarde se faisant sur les serveurs de la Ville. Les atouts essentiels du projet sont le portail

de coopération selon le protocole OAI et une gestion de contenu dynamique adaptée aux nouveaux usages du Web. Il est prévu notamment d'autoriser la correction d'OCR par les utilisateurs.

Projet de bibliothèque numérique et projet d'établissement sont étroitement liés dans une modernisation d'ensemble du réseau des médiathèques. A ce jour, certaines implantations n'offrent pas d'accès à Internet ni de collections numériques, les locaux et les équipements ne sont pas adaptés au numérique et aux usages nomades, ce qui a motivé un plan de remise à niveau de l'ensemble avec la création de médiathèques, le réaménagement de la bibliothèque centrale, la réinformatisation et la refonte du site web. L'ouverture du portail est prévue en octobre 2012.

Une convention de coopération numérique devrait être signée pour la période 2012-2014 entre la Ville d'Orléans, la Préfecture de région, la DRAC et la BnF, pour la numérisation et le signalement réciproque des documents dans Gallica. La bibliothèque numérique est aussi impliquée dans des projets de recherche, car elle coopère avec l'IRHT et participe à l'Equipex Biblissima pour reconstituer la bibliothèque virtuelle des manuscrits de Fleury. La BMVR se positionne aussi dans un rôle d'expert en Région Centre, et met l'accent sur les missions de coopération régionale. Les modalités techniques d'interopérabilité sont la clé de voûte du projet : le portail régional valorise les partenaires en préservant l'identité de chacun, accroît la visibilité des collections de chacun, et permet de réorienter les usagers vers chaque site pour l'accès direct aux documents.

Audition d'Albert Poirot et Frédéric Blin (Bibliothèque nationale et universitaire de Strasbourg)

Dans le secteur de l'enseignement supérieur, la Bibliothèque nationale et universitaire (BNU) de Strasbourg joue un rôle de premier plan en matière de numérisation. C'est la bibliothèque universitaire française la plus riche en volumes, avec trois millions de documents. Sa stratégie de numérisation s'appuie sur ses missions universitaires, régionales, nationales et européennes. Elle gère le dépôt légal imprimeur pour la région, elle assure la rédaction de la bibliographie alsacienne, et participe à Alsatica (portail des savoirs en Alsace). En tant que Centre d'acquisition et de diffusion de l'information scientifique et technique (CADIST) et pôle associé à la BnF, ses pôles d'excellence thématiques reconnus au plan national sont le domaine germanique et les sciences religieuses. La bibliothèque du Conseil de l'Europe lui a été transférée en 2009. Sa politique de numérisation se porte vers les fonds spécifiques, rares et précieux ainsi que vers les fonds sollicités par des programmes de recherche français ou internationaux, la numérisation systématique des fonds germaniques étant assurée par de grandes bibliothèques allemandes. La BNU s'appuie sur des partenariats avec la BnF, les universités (en premier lieu celle de Strasbourg, mais aussi Bordeaux dans le domaine juridique), la Staatsbibliothek de Berlin, le réseau EUCOR (Confédération des universités du Rhin supérieur), etc. Ces projets peuvent concerner des fonds numériquement restreints : une centaine de documents rares du XVI^e siècle sur les pays d'Europe du Nord, la participation à la reconstitution virtuelle de la Genizah du Caire, le projet Europeana Collections 1914-1918. Des numérisations sont aussi effectuées à l'occasion d'événements ou d'expositions de documents, à des fins de conservation ou pour la valorisation de supports spécifiques (monnaies, iconographie, papyrus, presse régionale). Mais les fonds numérisés (7 To) ne constituent pas encore une bibliothèque numérique car ils ne sont pas encore moissonnables, et la BNU s'apprête à lancer un appel à partenariat pour la diffusion de ses documents numériques.

Dotée d'un service de bibliothèque numérique comprenant cinq personnes, la BNU peut constituer pour les bibliothèques numériques de l'enseignement supérieur un des pôles d'expertise à la fois sur les aspects patrimoniaux et sur les aspects techniques. Du point de

vue de la BNU, pour que les bibliothèques numériques existent réellement en France, il conviendrait d'avancer sur deux points touchant à la législation : mettre l'accent sur les droits des usagers ; favoriser l'accès aux collections numérisées en application des directives gouvernementales Etalab avec la mise en place des licences d'exploitation de type *Creative commons*.

Audition de Nicole Bériou et Gilles Kagan (Institut de recherche et d'histoire des textes, CNRS)

La bibliothèque virtuelle des manuscrits médiévaux (BVMM) est emblématique d'une réalisation pour la recherche. C'est un entrepôt de consultation d'images mis en œuvre par l'IRHT, unité propre du CNRS dont les thèmes de recherche portent sur la globalité du manuscrit, de l'histoire du livre à celle des textes. L'IRHT, qui a constitué dès 1937 une collection de reproductions de manuscrits (microfilms, diapositives), mène depuis 1979, avec le soutien des ministères concernés des campagnes méthodiques de couverture des fonds patrimoniaux français des bibliothèques universitaires et municipales, la BnF ayant réservé ses droits pour son propre fonds. En 2008, l'IRHT est passé complètement du microfilmage à la numérisation haute définition en couleurs, selon des choix techniques pour la prise de vues conformes aux préconisations du MCC. Les métadonnées de description technique sont limitées au minimum requis, et les métadonnées textuelles sont très sommaires (signalement du manuscrit, nom du photographe, coordonnées de conservation,...). Pour les fonds des bibliothèques d'enseignement supérieur et de recherche, l'archivage pérenne est assuré par le CINES ; pour les fonds des bibliothèques de la Culture, l'IRHT est en pourparlers avec l'IN2P3 et la BnF. L'IRHT assure aussi dans ses campagnes locales une couverture photographique des fonds des archives et des musées qui conservent des manuscrits médiévaux, voire des incunables. Des conventions bipartites préservent un partage des droits, et les bibliothèques reçoivent une copie intégrale de la campagne photographique réalisée, chacun gardant les droits de diffusion qui lui sont propres. L'IRHT s'appuie à la fois sur sa collection de 75 000 microfilms, qui garde encore aujourd'hui un intérêt, et sur la numérisation intégrale de manuscrits. Elle intervient également dans des programmes spécifiques financés *ad hoc*, telle la numérisation, l'exploitation et la valorisation scientifique de la bibliothèque de Chartres sinistrée en 1944. La BVMM a trois sources d'accroissement : les campagnes de numérisation intégrale récentes en haute définition, les vues qui proviennent du décor des manuscrits, la numérisation des microfilms. Au total, on compte près de 650 000 vues.

La Fondation campus Condorcet a déposé, dans le cadre des investissements d'avenir un projet d'Equipex, *Biblissima*, en collaboration notamment avec l'IRHT, la BnF, l'Ecole pratique des hautes études, l'Ecole nationale des Chartes, l'université de Lyon pour son Centre d'histoire médiévale, l'université de Tours pour son Centre d'études supérieures sur la Renaissance. Le projet, sur une durée de 10 ans, porte sur environ 2 000 manuscrits médiévaux n'ayant jamais été traités et sur le retraitement des ouvrages insuffisamment ou incomplètement microfilmés. Cet Equipex associerait très étroitement à des programmes scientifiques les ressources méritant d'être traitées.

L'IRHT n'ayant qu'un droit de diffusion en interne, l'accès se fait uniquement dans sa bibliothèque sans possibilité pour le lecteur de reproduire ou de télécharger sur son ordinateur, et l'export est limité afin d'éviter la copie abusive. Il faudrait toutefois ouvrir plus largement la bibliothèque virtuelle¹. L'IRHT a en projet sur son site web un *atelier de lecture du livre médiéval*, qui permettrait de passer librement sur Internet du catalogue de manuscrits

¹ Le comité de pilotage de la BVMM qui s'est réuni le 1^{er} février 2012 a approuvé à l'unanimité la proposition d'ouvrir la BVMM sur Internet, après accord des bibliothèques partenaires.

enluminés « Initiale » à la BVMM. Pour cela, l'autorisation préalable des collectivités détentrices des manuscrits est nécessaire. Il est souhaitable de parvenir à une situation des droits d'exploitation qui favorise la circulation scientifique et la reproduction des images.

Audition de Marie-Luce Demonet (Centre d'Etudes Supérieures de la Renaissance)

Programme de numérisation à vocation patrimoniale et de recherche dans un cadre régional, les bibliothèques virtuelles humanistes (BVH) sont dans une démarche d'expérimentation continue. Les BVH sont mises en œuvre depuis 2003 dans le cadre de deux contrats Etat-Région successifs par le Centre d'études supérieures de la Renaissance (CESR), qui est à la fois une unité mixte de recherche CNRS et une composante de l'Université de Tours. Le CESR participe aussi à l'Equipex Biblissima, ainsi qu'au Labex Patrimar en partenariat avec les universités de Poitiers et de La Rochelle. L'objectif du CESR est de produire 2000 facsimilés d'ouvrages de la Renaissance en Région Centre (réalisé à 50%) et de transcrire en français une centaine de textes de la Renaissance. Les contenus des BVH, liés à des programmes de recherche, sont les classiques de la Renaissance : ouvrages majeurs et corpus de grands auteurs, sources religieuses, sources de la science, jurisprudence et textes politiques, ouvrages d'intérêt régional. Les publics des BVH sont en priorité des universitaires et des étudiants, mais aussi des amateurs, selon des objectifs variés de recherche, de formation et de culture.

Les campagnes de numérisation se sont déroulées au CESR (qui possède une bibliothèque spécialisée), ainsi que dans les BM de Tours, d'Orléans, de Blois, de Vendôme, de Châteauroux et de Bourges, et dans des bibliothèques privées. Les documents sont traités sous un double prisme, recherche et patrimonial. Le stockage est assuré par le CINES pour la conservation et les mises à jour, et par l'IN2P3 pour la diffusion. Une copie de sauvegarde étant également hébergée par la Maison des sciences de l'homme de Tours. Pour les traitements, le CESR privilégie les choix techniques standardisés avec une migration générale vers XML et XML/TEI, y compris pour les métadonnées. Pour les entrepôts OAI, le CESR recherche la meilleure compatibilité possible avec les différents moissonneurs : CCFR et Gallica (OAI-PMH) ; Europeana (TEI/MARC/Dublin Core). Par ailleurs, le centre travaille à de nouvelles fonctionnalités pour les facsimilés : feuilletage en simple et double page, pages en JPEG avec zoom et téléchargement page à page, nouvelle ergonomie de consultation et de présentation des PDF. Seront également proposés trois types de téléchargements : PDF couleurs sans feuilletage ni OCR, PDF en deux niveaux de qualité (sans OCR ; compressé avec OCR brut en texte caché). La recherche porte notamment sur un OCR permettant de différencier aussi bien la graisse et la taille des caractères que la police ou la fonte. Le coût moyen de numérisation d'un ouvrage de 250 pages est de 400€, qu'elle soit faite en interne ou externalisée.

Les BVH sont l'un des centres de ressources du TGE Adonis. Le CESR adhère au protocole des *Digital Humanities*, qui implique une ouverture maximale et un accès libre et gratuit à toutes les données. Les images issues des fonds publics sont une propriété partagée et le CESR exploite librement les images JPEG, les originaux étant gérés par les établissements d'origine. Le CESR travaille sous licence *Creative Commons*, en appliquant pour les images la clause la plus restrictive (pas de reproduction sans autorisation). Les BVH sont intégrées dans Europeana comme *data content provider*. Elles sont aussi intégrées dans les infrastructures DARIAH (*Digital Research Infrastructure for the Arts and Humanities*) pour les facsimilés et dans CLARIN pour la base textuelle. Pour l'avenir, les points d'attention sont notamment : la numérisation de qualité de petites unités qualitatives et sélectionnées ; l'intégration dans un ensemble plus vaste, le consortium CAHIER, corpus d'auteurs pour les humanités qui alimente la TGIR CORPUS.