BnF

Schéma numérique

Mars 2016

Sommaire

INTRODUCTION **03**
Les axes stratégiques de la BnF 05
Vision d'avenir 10

01. CONSTITUER UNE COLLECTION NUMERIQUE DE REFERENCE **14**

ZOOM SUR QUELQUES POINTS FORTS
01.A_ Le dépôt légal du web 22
01.B_ La numérisation des documents audiovisuels 23

MOYENS, OUTILS ET MÉTHODES
01.C_ La politique documentaire de la numérisation 24
01.D_ Les différents modes de financement de la numérisation 25
01.E_ Les ateliers internes de numérisation 26

AXES DE DÉVELOPPEMENT
01.F_ Le cas particulier de la presse 27
01.G_ La numérisation des œuvres encore soumises à la propriété intellectuelle 28
01.H_ Le dépôt légal des livres numériques 29

02. SIGNALER, CONSERVER **30**

ZOOM SUR QUELQUES POINTS FORTS
02.A_ Data.bnf.fr 36
02.B_ L'entrepôt d'archivage numérique SPAR 37

MOYENS, OUTILS ET MÉTHODES
02.C_ L'évolution du catalogue général vers le modèle FRBR 38
02.D_ Ressources électroniques acquises : signalement et accès 39
02.E_ Garantir une consultation pérenne des documents audiovisuels 40

AXES DE DÉVELOPPEMENT
02.F_ La transition bibliographique 41

03. INNOVER AU SERVICE DES PUBLICS **42**

ZOOM SUR QUELQUES POINTS FORTS
03.A_ Le site www.bnf.fr 50
03.B_ L'éditorialisation et la médiation des contenus de Gallica 51
03.C_ Un écosystème de communication numérique 52
03.D_ Les outils en ligne d'aide à la recherche 53
03.E_ Diffusion et valorisation commerciale des collections numériques 54
03.F_ Les publications scientifiques 55

MOYENS, OUTILS ET MÉTHODES
03.G_ Les profils et les attentes des publics de Gallica 56
03.H_ Les canaux de la politique culturelle numérique 58
03.I_ L'accessibilité numérique 59

AXES DE DÉVELOPPEMENT
03.J_ Dissémination des ressources de Gallica et stratégie des API 60
03.K_ Développer les usages grand public de Gallica 61
03.L_ L'éducation artistique et culturelle 62
03.M_ Le prêt numérique 63
03.N_ L'innovation pour la consultation de la presse 64
03.O_ Le déploiement du Wi-Fi 65

04. IRRIGUER UN RÉSEAU DE PARTENAIRES NATIONAUX ET INTERNATIONAUX **66**

ZOOM SUR QUELQUES POINTS FORTS
04.A_ Les programmes nationaux de numérisation concertée 74
04.B_ Le catalogue collectif de France (CCFr) 75
04.C_ L'offre de tiers archivage 76
04.D_ La coopération numérique internationale 77
04.E_ Europeana, la bibliothèque numérique du patrimoine culturel européen 78

MOYENS, OUTILS ET MÉTHODES
04.F_ Gallica marque blanche 79
04.G_ L'action dans les instances internationales 80
04.H_ La rubrique « Pour les professionnels » du site institutionnel 81

AXES DE DÉVELOPPEMENT
04.I_ L'accès en région aux archives de l'internet 82
04.J_ La diffusion des documents numérisés par d'autres institutions patrimoniales 83

05. SIMPLIFIER LES OUTILS ET LES PROCESSUS POUR LA BNF ET SES PARTENAIRES **84**

ZOOM SUR QUELQUES POINTS FORTS
05.A_ La chaîne de numérisation : processus et système 92
05.B_ L'extranet du dépôt légal 93
05.C_ L'extranet Espace Coopération 94
05.D_ La plateforme PLATON pour l'exception handicape 95

MOYENS, OUTILS ET MÉTHODES
05.E_ La gestion électronique des documents 96
05.F_ Les outils de l'archivage du web 97
05.G_ La qualité 98

AXES DE DÉVELOPPEMENT
05.H_ Les outils de gestion des collections numériques 99
05.I_ Les acquisitions et dons de collections numériques (ADDN) 100

06. ENGAGER UNE DEMARCHE D'INNOVATION DURABLE **102**

ZOOM SUR QUELQUES POINTS FORTS
06.A_ De nouveaux formats pour le livre numérique : EPUB, DAISY 108
06.B_ Recherche et développement autour de l'océrisation 109

MOYENS, OUTILS ET MÉTHODES
06.C_ Les moyens humains 110
06.D_ L'investissement dans le développement informatique 112
06.E_ Communication interne et accompagnement au changement 113
06.F_ La formation continue 114

AXES DE DÉVELOPPEMENT
06.G_ La fouille de textes et de données (*text & data mining*) 115
06.H_ La numérisation en 3D 116
06.I_ Le projet de coopération avec le CEA 117

Introduction

En moins de deux décennies, la révolution numérique a provoqué des bouleversements majeurs, qui représentent un défi perpétuellement renouvelé pour l'action publique. Dans le domaine culturel, les efforts entrepris en matière de développement numérique ont accompagné très tôt ces évolutions. Socle de la mémoire des hommes et de l'histoire collective, le patrimoine s'est affirmé comme lieu d'innovations fécondes, inédites par leur ampleur et leurs implications.

Les technologies numériques, par la transformation profonde des outils, des objets et des usages qu'elles impliquent, affectent directement les missions d'une bibliothèque nationale : conserver, collecter, enrichir, transmettre. Elles éclairent une réflexion plus vaste, liée à la continuité et à l'adaptabilité d'un service public qui évolue sans cesse, au nom de la démocratisation culturelle, de l'égalité d'accès et, au-delà même de la gratuité, de l'appropriation par tous des nouveaux contenus, supports et outils numériques. Un défi qui se décline en termes de médiation et d'action culturelle, de diversification des publics, mais également d'accessibilité.

Comment penser l'avenir du patrimoine historique et culturel national à l'heure où s'élaborent les fondations d'une ère nouvelle ? Dans quelle mesure, et pour quel(s) usage(s), reconnaître le numérique comme outil et objet patrimonial ? Comment créer une continuité des âges alors même que les contenus, les supports, les grandeurs ne cessent d'évoluer ? Comment faire tenir ensemble formes patrimoniales traditionnelles et nouvelles, largement en devenir ?

En vertu de ses missions statutaires, la Bibliothèque nationale de France (BnF) a toujours veillé à répondre à ces défis. En 1988, alors que s'élaborait le projet de grande bibliothèque, Jacques Attali envisageait que la bibliothèque « d'un genre entièrement nouveau » voulue par François Mitterrand fût entièrement dématérialisée. L'idée a pris corps par étapes : premières numérisations, naissance de la bibliothèque numérique Gallica, début de la numérisation de masse, conversion de l'organisation au numérique…

Des rapports importants remis aux ministres de la Culture et de la Communication ont scandé ces années de questionnement et de déploiement. Ils avaient pour thème *La numérisation du patrimoine écrit* (Marc Tessier, 2010), le *Schéma numérique des Bibliothèques* (Bruno Racine, 2010) ou encore les *Contributions aux politiques culturelles à l'ère numérique* (Pierre Lescure, 2013). Tous à leur manière disaient la centralité nouvelle de la question numérique – pour les pouvoirs publics, le législateur, la BnF et les citoyens.

Texte cadre associé à une carte heuristique, ce *Schéma numérique* de la BnF rend compte des réflexions et des actions qui, en l'espace de quelques années, ont conféré à la Bibliothèque un rôle pionnier dans l'actualisation des politiques patrimoniales, en France comme à l'étranger.

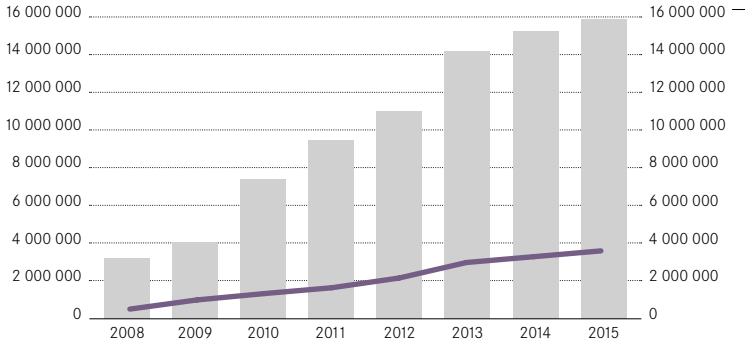
L'entrée des bibliothèques dans l'ère numérique soulève de nombreux défis que l'établissement analyse, intègre et, à maints égards, devance. Ce travail à la fois pragmatique et prospectif trace le chemin qu'empruntera la BnF dans les années à venir, tâchant d'offrir un service en phase avec les attentes et les usages actuels, tout en anticipant et en participant aux prochaines évolutions d'un environnement technologique en perpétuelle mutation.

Le Schéma numérique de la BnF rend compte des réflexions et des actions qui ont conféré à la Bibliothèque un rôle pionnier dans l'actualisation des politiques patrimoniales, en France comme à l'étranger.

Cette stratégie numérique ne saurait être mise en œuvre sans une politique ambitieuse et volontariste en matière d'accroissement des collections, de conservation, de signalement, d'innovation, de médiation, de coopération, de formation. Il en va de la capacité de l'institution à préserver un patrimoine commun, à le développer, à l'enrichir et à en favoriser la diffusion, au bénéfice du plus grand nombre.

Ces enjeux, la Bibliothèque les a traduits et articulés en six axes stratégiques prioritaires. Ils forment la trame du *Schéma numérique*. En voici l'exposé synthétique.

GALLICA
NOMBRE DE VISITES



GALLICA
NOMBRE DE DOCUMENTS

01
CONSTITUER UNE COLLECTION NUMÉRIQUE DE RÉFÉRENCE

La constitution d'une collection numérique de référence recouvre trois grands types d'opérations.

La première d'entre elles est de numériser des documents existant en format traditionnel : ouvrages et journaux, bandes vidéo, photographies argentiques, médailles… Touchant des documents de natures très diverses, cette numérisation dite « rétrospective » privilégie les livres et périodiques entrés dans le domaine public. Épuisées ou non, des œuvres sous droits peuvent aussi être numérisées en raison de leur intérêt patrimonial, de leur pertinence scientifique ou de leur fragilité matérielle.

La seconde est de collecter des œuvres existant déjà en format numérique : les documents nativement numériques entrant par dépôt légal, les ressources électroniques acquises (livres numériques, revues, bases de données, documents audiovisuels, etc.), les dons…

La dernière de ces opérations consiste à archiver le web. Mis en place en août 2006, le dépôt légal du web joue un rôle prépondérant dans l'accroissement des collections numériques de la BnF. Des campagnes d'archivage du web national, larges (noms de domaine en « .fr » notamment) ou plus ciblées (élections, presse, autres collectes thématiques) sont régulièrement menées, consistant à collecter en ligne sites, pages et fichiers.

Toutes ces formes d'enrichissement des collections visent à conjuguer volume et encyclopédisme, au profit de la communauté scientifique, éducative, ou du grand public.

Toutes ces formes d'enrichissement des collections poursuivent un objectif d'exhaustivité et, là où l'abondance des contenus en ligne suppose d'être sélectif, visent à conjuguer volume et encyclopédisme, au profit de la communauté scientifique, éducative, ou du grand public.

02
SIGNALER ET CONSERVER

La constitution d'un fonds patrimonial de référence a deux corollaires : garantir la conservation durable des documents et favoriser leur signalement, c'est-à-dire leur faculté à être identifiés, décrits et trouvés. Il s'agit là de deux expertises historiques de la BnF. À l'heure d'internet et de l'ouverture des données, la BnF doit développer et parfois réinventer ces expertises en les appliquant à de nouveaux objets, de nouveaux supports, de nouveaux usages.

La constitution d'un fonds patrimonial de référence a deux corollaires : garantir la conservation durable des documents et favoriser leur signalement.

Pour faire face à l'évolutivité des formats et à l'obsolescence des supports numériques, la BnF a ainsi mis en place un système de conservation perfectionné, SPAR (Système de préservation et d'archivage réparti). Véritable entrepôt numérique, SPAR offre les meilleures garanties de préservation et de sécurité des données. Agréé par le SIAF (Service interministériel des Archives de France), le système effectue de multiples copies des documents archivés et contrôle régulièrement leur intégrité. Il contribue aussi à identifier les formats technologiquement obsolètes, permettant leur transformation.

Garantir la pérennité des documents est une chose ; les rendre visibles en est une autre, d'où l'enjeu décisif de l'ouverture et de la dissémination des métadonnées. Pour permettre aux informations bibliographiques de circuler le plus possible sous forme numérique, un ambitieux travail est mené depuis plusieurs années par les équipes de la BnF autour du signalement. Travail sur les normes de catalogage, restructuration des données, développement du web sémantique, convergence des catalogues… : les chantiers ne manquent pas pour garantir le meilleur référencement possible des données dans les moteurs de recherche.

Le lancement du service data.bnf.fr est emblématique de cette volonté d'améliorer l'accès aux métadonnées bibliographiques de la BnF. Depuis janvier 2014, celles-ci sont placées sous licence ouverte de l'Etat. L'audience de la Bibliothèque sur le web s'en trouve renforcée, confortant la diffusion de la langue et de la culture françaises à travers le monde.

03
INNOVER AU SERVICE DES PUBLICS

Offrir au plus grand nombre la possibilité d'accéder massivement et durablement aux ressources de la création et du savoir, telle est la responsabilité première d'une bibliothèque, a fortiori lorsque sa vocation est universelle.

Clé de voûte de la collection numérique de la BnF, la plateforme Gallica est l'instrument premier de cette mission. Avec plus de 3 millions de documents, elle figure parmi les premières bibliothèques numériques au monde. Par son intermédiaire, une partie importante des fonds de la Bibliothèque est diffusée et valorisée en ligne. Un tel accès élargi aux ressources impose de relever de nombreux défis.

La BnF a mis en œuvre une politique d'hospitalité numérique à destination de ses publics sur place.

Le premier des défis consiste à éviter que le nombre considérable de données, de fonctionnalités, d'applications nuise au confort et à l'efficacité des recherches. Ce n'est plus le silence qui pose aujourd'hui problème, mais le bruit. La qualification des métadonnées, la simplification des interfaces, l'optimisation pour les moteurs de recherche, le développement de portails thématiques, la création de services d'aide en ligne sont autant de manières pertinentes de tirer profit des possibilités offertes par les nouvelles technologies.

Les publics de la BnF étant par nature pluriels, il est important de proposer des services adaptés à la singularité des besoins des communautés d'utilisateurs : chercheurs, étudiants, professionnels, particuliers, publics éloignés de la culture… La conception de programmes éditoriaux dans les champs littéraires, scientifiques, historiques, à partir des fonds et documents détenus par la Bibliothèque s'élabore en correspondance avec cette diversité des publics, de même que sont instaurés des niveaux d'accès différenciés, permettant aux lecteurs des salles de recherche de consulter des documents sous droits via les versions intra muros de Gallica.

Enfin, innover au service des publics suppose de prendre acte des usages de l'*homo numericus*. La BnF a mis en œuvre une politique d'hospitalité numérique à destination de ses publics sur place : Wi-Fi, possibilité de photographier les documents, services en accès direct sur les ordinateurs personnels et en mobilité (réservation de places et de documents, jauge des salles, consultations de ressources numériques, commande en ligne de travaux de reproduction, espaces personnels…). Les nouvelles pratiques de recherche (veille et partage sur les réseaux sociaux, administration de blogues et de sites internet, constitution de corpus numériques) disposent en outre la communauté scientifique à devenir actrice à part entière de la politique documentaire de la Bibliothèque. À terme, une partie des données pourraient être coproduites avec les chercheurs, grâce au développement d'outils d'enrichissement de contenus, de correction, de renvoi, selon une logique de service. Plus généralement, en favorisant la constitution de communautés d'usagers, tels les Gallicanautes, en se montrant très présente sur les réseaux sociaux (Facebook, Twitter, Pinterest…) et en facilitant la reprise de ses contenus sur d'autres sites, la BnF se donne les moyens d'accroître la visibilité de ses services et la circulation de ses ressources.

À terme, une partie des données pourraient être coproduites, grâce au développement d'outils d'enrichissement de contenus, de correction, de renvoi, selon une logique de service.

La richesse de leurs fonctionnalités et la perspective de leurs développements futurs font de Gallica, Sindbad (service d'aide en ligne ou par téléphone) ou encore AVEC (portail d'accès aux ressources numériques sur les terminaux personnels des usagers) de puissants vecteurs d'innovation et d'amélioration des services rendus au public. Ils contribuent à une plus grande mobilité et à une plus grande personnalisation des usages.

04
IRRIGUER UN RÉSEAU DE PARTENAIRES NATIONAUX ET INTERNATIONAUX

Du fait de son histoire, de la richesse de ses collections, de son positionnement institutionnel, mais aussi de sa mission de coopération, la BnF joue un rôle moteur en matière de politique numérique.

Ceci est vrai à l'échelle nationale, où elle irrigue un réseau de partenaires comprenant des bibliothèques en région, des institutions de recherche, des établissements culturels. Consistant tour à tour en opérations de numérisation, de signalement et de mise en valeur, ces actions de coopération répondent à une exigence de redistribution des moyens financiers alloués chaque année par les pouvoirs publics – le Centre national du livre en premier lieu –, au titre de la numérisation du patrimoine culturel français. En retour, ces rapprochements permettent à la BnF d'enrichir les collections numériques nationales et de renforcer leur visibilité.

L'offre « Gallica marque blanche » illustre parfaitement l'effet d'entraînement que les projets de la BnF sont en mesure de provoquer en matière de numérisation.

En offrant la possibilité à un partenaire de disposer de sa propre bibliothèque numérique (en contrepartie de quoi le partenaire autorise l'intégration d'une copie de ses ressources numériques dans les collections nationales), l'offre « Gallica marque blanche » illustre parfaitement l'effet d'entraînement que les projets de la BnF sont en mesure de provoquer en matière de numérisation.

L'expertise et la créativité des équipes de la BnF sont aussi mises à profit au niveau international. Elle joue ainsi un rôle décisif dans le développement d'Europeana, la bibliothèque numérique du patrimoine historique et culturel de l'Europe. Elle exerce une influence majeure dans les instances internationales qui élaborent normes et standards, établissent des programmes de coopération numérique ou développent de nouveaux outils : Réseau francophone numérique, IIPC (*International Internet Preservation Consortium*), IFLA (*International Federation of Library Associations*), LIBER (Ligue des bibliothèques européennes de recherche), Fondation Europeana, CENL (Conférence des bibliothèques nationales européennes), et dernièrement le consortium IIIF (*International Image Interoperability Framework*).

Des partenariats sont noués auprès de bibliothèques, d'universités, d'établissements culturels, sur les cinq continents. En résultent des programmes de numérisation partagée, des transferts de compétences, le développement de portails thématiques… Ces partenariats valorisent l'expertise française en matière de conservation et de diffusion patrimoniale. Ils sont de puissants leviers de rayonnement pour notre pays, sa culture et sa langue.

La BnF exerce une influence majeure dans les instances internationales qui élaborent normes et standards, établissent des programmes de coopération numérique ou développent de nouveaux outils.

05
SIMPLIFIER LES OUTILS ET LES PROCESSUS POUR LA BNF ET SES PARTENAIRES

Bibliothèques, institutions culturelles, professionnels des industries culturelles, associations : la BnF se situe à la confluence de pratiques et de besoins dont la densité, l'intensité et la diversité se sont accrues sous l'effet des technologies numériques. Une simplification des outils et des processus s'impose donc pour fluidifier l'écosystème dans lequel la Bibliothèque prend place.

La simplification a été engagée au niveau des relations avec les partenaires, dans le sens d'une meilleure mutualisation des outils et d'une plus grande autonomie des acteurs. Un « Espace Coopération » a ainsi été institué en 2011. Facilitant les opérations partenariales de suivi des collections référencées ou numérisées par la BnF, il permet également le versement de documents numériques dans Gallica ou l'accès aux données pour les bénéficiaires du tiers archivage. Il permettra sous peu le paramétrage des bibliothèques numériques en Gallica marque blanche.

La simplification a été engagée au niveau des relations avec les partenaires (bibliothèques, institutions culturelles, professionnels des industries culturelles, associations), dans le sens d'une meilleure mutualisation des outils et d'une plus grande autonomie des acteurs.

Les procédures de dépôt légal ont elles aussi été simplifiées, via la mise en place d'un service extranet de déclaration en ligne. Les données bibliographiques sont saisies par les déposants avant d'être récupérées dans le système d'information de la BnF, les éditeurs devenant de la sorte coproducteurs des métadonnées qui sont ensuite enrichies par les catalogueurs.

Un important effort de dématérialisation des procédures a été entrepris, doublé du déploiement d'un dispositif de gestion électronique des documents (GED).

Mis en place en 2010, le service d'édition adaptée pour les publics en situation de handicap (plateforme PLATON) témoigne également de l'évolution des relations partenariales vers plus d'autonomie et plus de mutualisation : les fichiers numériques ayant servi à l'édition d'une œuvre sont déposés par l'éditeur auprès de la BnF, qui joue dès lors le rôle de tiers de confiance pour assurer la transmission de ces fichiers aux organismes habilités qui en ont fait la demande et réalisent l'édition adaptée (braille, gros caractères…) de ces ouvrages.

L'impératif de simplification et d'optimisation des outils et procédures vaut aussi pour la BnF, s'agissant de la gestion des échanges et organisations internes aussi bien que de celle des collections numériques. Un important effort de dématérialisation des procédures a ainsi été entrepris, doublé du déploiement d'un dispositif de gestion électronique des documents (GED) régulièrement actualisé. La mise en place de ce dernier s'est accompagnée d'un important travail de formation et d'adaptation, au profit d'un fonctionnement modernisé et fluidifié de l'organisation. Des instruments de pilotage interne et des indicateurs fiables gouvernent la relation contractuelle avec les prestataires, dans le cadre des marchés de numérisation, notamment.

06
ENGAGER UNE DÉMARCHE D'INNOVATION DURABLE

La BnF ne serait pas en mesure d'accompagner efficacement les changements induits par le numérique, ni a fortiori de les anticiper, si elle n'était pas elle-même un important acteur de l'innovation et de la recherche sur le plan des usages et des technologies. Tout en mobilisant des moyens significatifs tant en ressources humaines qu'en développement informatique, la Bibliothèque se dote de dispositifs qui favorisent l'accompagnement au changement et assurent sa capacité à continuer d'évoluer dans un environnement en mutation permanente.

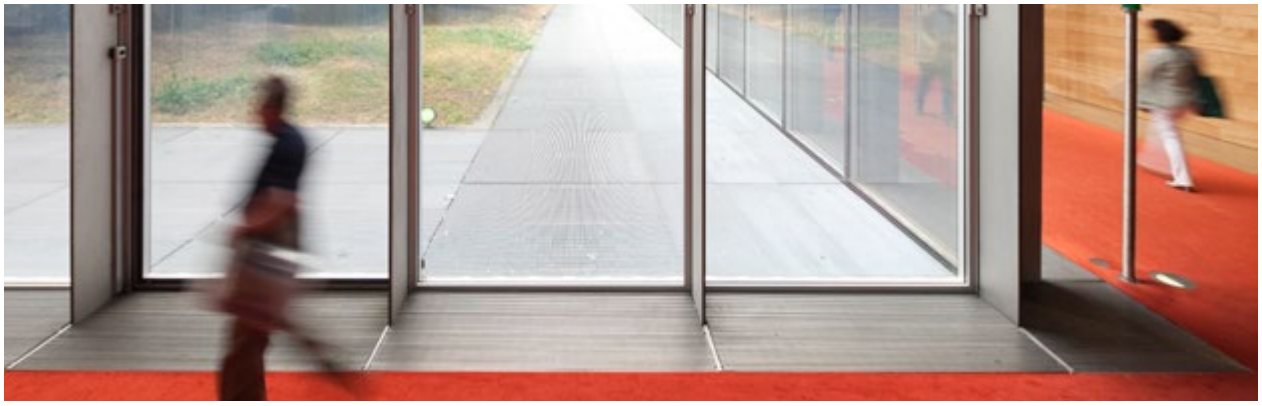
La Bibliothèque est partenaire de plusieurs laboratoires d'excellence et a signé des accords de coopération avec des établissements de recherche et d'enseignement supérieur.

D'une certaine manière, l'innovation a partie liée avec l'ensemble des initiatives prises par l'établissement pour améliorer concrètement la conservation et la diffusion de ses richesses patrimoniales. Si l'on se place du point de vue des usagers et de leurs besoins, l'innovation consiste principalement à élargir les possibilités de recherche, de consultation, de croisement des données relatives à un nombre toujours croissant de documents. L'instauration d'un système de prêt numérique ou l'amélioration des paramètres d'interrogation et de navigation textuelle des fonds de presse pourraient être quelques-unes des innovations majeures introduites par la BnF dans un avenir proche.

En matière de recherche aussi l'établissement fait montre d'une intense activité. La Bibliothèque est partenaire de plusieurs laboratoires d'excellence et a signé des accords ponctuels de coopération avec des établissements de recherche et d'enseignement supérieur. Ces partenariats portent entre autres sur les usages du patrimoine numérique, la fouille de texte et de données, l'éditorialisation ou l'exploration de fonds ciblés, contribuant de ce fait au développement des humanités numériques.

Les ingénieurs et techniciens de la BnF conduisent, par ailleurs, d'importantes expérimentations techniques, en lien le plus souvent avec des partenaires publics ou privés. Ces travaux portent notamment sur les nouveaux systèmes de production de l'information numérique, la production participative de contenus, l'amélioration de la reconnaissance optique de caractères (OCR).

Autant de développements susceptibles de rencontrer demain les besoins des utilisateurs de la Bibliothèque ; autant de programmes, aussi, par le biais desquels les équipes de la BnF et leurs partenaires de recherche et de coopération apportent une contribution significative à l'affirmation de l'excellence française en matière de numérique.



Vision d'avenir

La cartographie du numérique à la BnF dessinée dans le présent document, en faisant émerger pour chaque axe stratégique des points forts, des moyens et des axes de développement, permet d'esquisser une vision prospective à moyen terme. Si l'effort d'innovation et de changement se poursuivait au même rythme et avec sensiblement les mêmes moyens, à quoi ressemblerait la BnF numérique dans cinq ans ? Quels seraient les jalons atteints, les principaux caps franchis ? Sans vouloir anticiper le travail de projection des actions et des objectifs qui s'imposera avec le nouveau contrat de performance à partir de 2017, nous proposons ici quelques pistes de réflexion qui contribuent à faire apparaître la silhouette de la Bibliothèque de 2020.

La BnF, lieu d'hospitalité numérique

Le lecteur accueilli à la Bibliothèque, aussi bien sur le site François-Mitterrand qu'à Richelieu, bénéficiera d'outils attractifs : des ressources numériques exclusives mises en valeur de façon claire, un site web renouvelé pour mieux répondre à ses besoins, des services accessibles et utilisables en mobilité. Il se verra proposer des services personnalisés, de l'accès au Wi-Fi à la fouille de corpus numériques. La BnF saura également surprendre ses visiteurs, notamment à Richelieu, en leur offrant une expérience augmentée du patrimoine, source renouvelée de sensations et d'émotions. La BnF utilisera ainsi le numérique comme levier pour dynamiser sa fréquentation sur place. En ligne, elle conviera les internautes à de nouvelles explorations virtuelles traversant l'histoire de ses sites et de ses collections : une stratégie de diffusion culturelle numérique susceptible d'éveiller la curiosité des publics jeunes et de renforcer ses relais dans le domaine pédagogique comme dans celui du champ social.

Gallica, source de culture et creuset de tous les publics

Chercheurs, étudiants, amateurs mais aussi généalogistes, passionnés, curieux, enfants, familles : les Gallicanautes seront plus nombreux et leurs usages plus diversifiés grâce à une stratégie de médiation ciblée en direction du grand public, tant sur les sites de la BnF que sur les réseaux sociaux ou via des applications mobiles. La recherche de dispositifs nouveaux confortera la BnF dans son rôle d'établissement explorant de nouveaux horizons pour l'accès de tous à la culture : prêt numérique, animation de communautés virtuelles de contributeurs (Gallica studio).

Gallica et Gallica intra muros, creusets de tous les savoirs

La richesse toujours croissante des contenus proposés sur Gallica et la qualité des sélections documentaires contribueront à l'attractivité de la bibliothèque numérique. Sur Gallica seront consultables les numérisations les plus qualitatives d'un point de vue intellectuel et technique, afin d'améliorer l'expérience de l'internaute en lui offrant une navigation claire et précise. Des extraits de documents sous droits lui seront également proposés. Sur Gallica intra muros, les lecteurs fréquentant les sites physiques de la BnF se verront réserver l'exhaustivité de la collection patrimoniale numérisée, qui passera la barre des 5 millions de documents grâce à la poursuite de l'effort de numérisation national. 500 000 fascicules de presse supplémentaires et 200 000 livres indisponibles du vingtième siècle numérisés dans le cadre du projet ReLIRE y côtoieront plusieurs dizaines de milliers de livres numériques entrés par dépôt légal, des collections spécialisées de plus en plus volumineuses, les numérisations réalisées en partenariat avec des acteurs privés et les collections numériques des partenaires membres du réseau national de coopération.

La patrimonialisation du numérique en marche

La BnF sera pleinement adaptée, aussi bien sur le plan juridique que technique, aux nouvelles modalités de circulation des biens culturels et autres contenus numériques destinés à enrichir ses collections. En développant le dépôt légal du livre numérique et en automatisant celui des documents sonores, mais aussi celui des vidéos et d'autres documents web et multimédias qui lui échappent encore aujourd'hui, elle aura fait prendre à sa collection patrimoniale un virage décisif vers le numérique, reflétant l'évolution de la société. Grâce à des outils et des processus sécurisés et performants, elle sera identifiée comme un partenaire de confiance par les acteurs de la chaîne de l'édition, avec qui elle échangera données et contenus de manière fluide. Elle accueillera, sous forme de dons ou d'achats, des documents numériques de toute forme ou nature, dans le prolongement de ses collections traditionnelles aussi bien que dans l'exploration de formes nouvelles.

L'excellence de la préservation numérique

La BnF continuera à développer sa capacité de préservation numérique et intégrera l'ensemble de ses filières dans SPAR, le Système de préservation et d'archivage réparti. Tous les fichiers issus de la numérisation, du dépôt légal du web et du livre numérique, ainsi que de l'archivage administratif de l'établissement seront préservés sur deux sites, tandis qu'un troisième, sécurisé pour faire face aux risques les plus aigus, sera recherché afin d'accueillir au moins les collections les plus sensibles. La convergence avec le système dédié à la préservation des documents audiovisuels sera mise en chantier. Enfin, la BnF continuera de partager son savoir-faire et son infrastructure dans ce domaine avec d'autres acteurs publics, notamment dans le cadre du tiers archivage.

Vers l'appropriation des mégadonnées

La BnF capitalisera sur l'expérience de data.bnf.fr pour faire évoluer en profondeur les modalités de production de ses métadonnées. Celle-ci sera plus collaborative : la production des référentiels et données d'autorité se fera au niveau national, au moins en partie, grâce à des outils partagés et à des standards comme l'ISNI qui rapproche la communauté des archives, des musées, des bibliothèques, de l'enseignement supérieur et de la recherche, et de l'édition. Les algorithmes, les flux de données et d'une façon plus générale les outils du *big data* (mégadonnées) permettront à la Bibliothèque d'être plus efficiente, mais aussi d'apprendre à croiser les données bibliographiques avec d'autres données (traces d'usages, données de gestion…) afin de piloter l'évolution de ses outils et de ses services.

En développant le dépôt légal du livre numérique et en automatisant celui des documents sonores, la BnF fera prendre à sa collection patrimoniale un virage décisif vers le numérique.

La valeur de l'ouverture des données

Le savoir-faire de la BnF dans le domaine des mégadonnées favorisera également l'accès à ses ressources et leur visibilité. Dans le catalogue général, comme dans Gallica, les internautes et lecteurs trouveront directement les œuvres, les personnes, les lieux et les concepts qu'ils recherchent, grâce aux regroupements sémantiques et aux expériences de visualisation testés dans data.bnf.fr. Ce dernier s'ouvrira à davantage de données extérieures, à la fois pour les valoriser et pour mieux satisfaire l'internaute dans ses recherches. La prolifération des liens entrants et sortants, via des plateformes comme Wikipédia et Wikidata, inscrira résolument la BnF dans l'écosystème du web. Grâce à l'ouverture juridique des métadonnées depuis 2014, puis à l'ouverture technique via des API telles que IIIF et d'autres mécanismes de dissémination, les réutilisations des données de la BnF par des tiers prospéreront, ce qui contribuera à asseoir son image et sa notoriété. Cet élan bénéficiera aussi à ses partenaires, grâce à l'ouverture des données du Catalogue Collectif de France et au rôle d'agrégateur joué par la BnF pour la bibliothèque numérique Europeana.

Gallica, une infrastructure mutualisée

Gallica ne sera plus seulement la bibliothèque numérique de la BnF et de ses partenaires, mais le cœur d'une véritable infrastructure mutualisée pour un réseau toujours plus vaste d'acteurs publics, aussi bien en France qu'à l'international et notamment dans le monde francophone. Grâce au dispositif Gallica marque blanche, qui aura évolué pour devenir plus fluide et plus simple, mais aussi à d'autres outils de coopération numérique comme les portails de médiation internationaux « patrimoines partagés », la BnF apportera sa contribution au rayonnement numérique de ses partenaires. Des collections numérisées non encore accessibles faute d'une interface de consultation deviendront visibles pour le public. En plus d'accroître ses collections disponibles sur place et à distance, la BnF aura œuvré en faveur de l'accès de tous à la connaissance, en partageant avec ses réseaux de partenaires aussi bien son savoir-faire que les investissements consentis depuis plusieurs années.

La maîtrise des outils professionnels

Les systèmes d'information de la BnF continueront d'évoluer pour mieux accompagner la transition numérique. Les professionnels de la Bibliothèque seront dotés d'outils de pilotage et de gestion pour une meilleure maîtrise des collections aussi bien physiques que numériques, en veillant au nécessaire continuum entre les supports à des fins de conservation et d'accès. À travers ses extranets, la BnF étendra cette maîtrise à ses partenaires qui pourront agir de manière autonome sur l'entrée et l'accès de leurs documents et métadonnées. L'amélioration des procédures internes bénéficiera aussi bien aux lecteurs, à travers des dispositifs comme la numérisation en urgence dans les ateliers internes, qu'aux agents de la BnF qui pourront aiguiller les documents numériques vers l'environnement de publication approprié, ou encore aux professionnels de l'édition qui se feront les relais des ressources iconographiques valorisées dans la nouvelle banque d'images.

La BnF, laboratoire d'innovation

Enfin, la BnF continuera d'innover, d'être moteur de changement et d'inventer de nouvelles manières de conduire ses missions. Dans le sillage de sa collaboration avec l'Ideas Laboratory du CEA, elle se sera dotée d'une organisation dédiée, de dispositifs exploratoires et de méthodologies appropriées favorisant l'émergence de nouveaux outils et de nouveaux services. Acteur de la recherche, notamment dans le domaine des humanités numériques, elle s'ouvrira également à l'expérience d'autres métiers, y compris au monde de l'entreprise, en bénéficiant de la proximité de la Halle Freyssinet. Elle adoptera des formes d'accompagnement au changement et des modes de travail collaboratifs qui permettront à chacun de ses agents de s'approprier ces nouveautés et de contribuer à les imaginer et à les construire.

En plus de faciliter ses missions premières de collecte, de conservation et de catalogage, le numérique permet à la BnF de favoriser l'accès, la diffusion et la mise en valeur de ses collections, de servir la recherche scientifique, ou encore de renforcer la coopération entre établissements culturels. Le présent document s'efforce de montrer comment la Bibliothèque nationale de France a su profiter de la révolution numérique comme d'une formidable opportunité.

C'est en relevant ces nouveaux défis que la BnF, fidèle à ses missions historiques, continue d'assumer les responsabilités qu'implique la transmission de notre patrimoine.

01 CONSTITUER UNE COLLECTION NUMÉRIQUE DE RÉFÉRENCE - P14

02 SIGNALER ET CONSERVER - P30

03 INNOVER AU SERVICE DES PUBLICS - P42

04 IRRIGUER UN RÉSEAU DE PARTENAIRES NATIONAUX ET INTERNATIONAUX - P66

05 SIMPLIFIER LES OUTILS ET LES PROCESSUS POUR LA BNF ET SES PARTENAIRES - P84

06 ENGAGER UNE DÉMARCHE D'INNOVATION DURABLE - P102

{BnF

01

CONSTITUER UNE COLLECTION NUMÉRIQUE DE RÉFÉRENCE

3,5 millions
de documents libres de droits en ligne

1 million
de documents sous droits numérisés
dont plus de 700 000 audiovisuels

1 519 téraoctets
de documents dans Gallica

668 téraoctets
d'archives du web

305
partenaires de numérisation

15,7 millions
de vues numérisées en 2015

À l'heure où le public s'attend à trouver sur la toile, en un clic, la totalité de la production éditoriale passée et présente, le numérique est un enjeu stratégique pour la diffusion des savoirs et la démocratisation de la culture. Pour une bibliothèque nationale, il se traduit par la nécessité de constituer une collection patrimoniale numérique de référence, sans discontinuité entre les documents nativement numériques et les fonds plus anciens, afin que l'ensemble de ces ressources soient facilement accessibles à tous, aujourd'hui comme demain.

ZOOM SUR QUELQUES POINTS FORTS

01.A_ Le dépôt légal du web
01.B_ La numérisation des documents audiovisuels

MOYENS, OUTILS ET MÉTHODES

01.C_ La politique documentaire de la numérisation
01.D_ Les différents modes de financement de la numérisation
01.E_ Les ateliers internes de numérisation

AXES DE DÉVELOPPEMENT

01.F_ Le cas particulier de la presse
01.G_ La numérisation des œuvres encore soumises à la propriété intellectuelle
01.H_ Le dépôt légal des livres numériques

La numérisation des collections

La numérisation du patrimoine écrit, audiovisuel et graphique a pour ambition de couvrir peu à peu l'ensemble des époques, des types de documents, des disciplines, et d'y donner accès à travers Gallica. Aujourd'hui, grâce aux chantiers de numérisation de masse commencés au milieu des années 2000, 3,5 millions de documents sont en ligne dans Gallica. L'accroissement se poursuit régulièrement à raison de plusieurs millions de pages ou d'images par an, toutes filières confondues – imprimés, collections spécialisées, etc.

Trois grandes périodes se détachent dans la poursuite de cette entreprise, correspondant aux évolutions techniques et aux orientations de la politique documentaire :

2002-2005 : constitution d'une **bibliothèque de l'honnête homme**, autour des œuvres les plus significatives, discipline par discipline, de l'Antiquité à nos jours. La numérisation se faisait principalement en mode image.

L'accroissement de la numérisation se poursuit régulièrement à raison de plusieurs millions de pages ou d'images par an, toutes filières confondues.

2005-2009 : émergence de la **numérisation de masse**, en réaction à l'initiative de Google. Dès 2007, la reconnaissance optique de caractères (OCR) devient systématique et est appliquée rétrospectivement aux fonds déjà numérisés. À l'inverse de la période précédente, la numérisation de masse procède par désélection : une fois le fonds choisi, on en retire les documents protégés au titre de la propriété intellectuelle, les documents étrangers et ceux trop fragiles pour supporter les opérations de numérisation. L'histoire de France (26 %) et la littérature (25 %) sont les disciplines les plus représentées quantitativement.

Depuis 2009 : une numérisation plus collective, plus diversifiée dans ses contenus et ses financements. La numérisation de masse se poursuit tout en s'efforçant d'équilibrer les objectifs de quantité et de sélectivité ; commencée en 1998, la numérisation des documents audiovisuels fait son apparition dans Gallica.

La priorité est ainsi donnée aux documents que la BnF est seule à conserver. La Bibliothèque s'appuie, par ailleurs, de plus en plus sur une démarche de partenariats avec des établissements français et étrangers et ouvre ses chaînes de numérisation aux collections d'autres institutions. Des dispositifs sont mis en œuvre pour ne pas limiter cette entreprise de numérisation aux seules œuvres du domaine public, ce qui obligerait à exclure les publications de la plus grande partie du XXe siècle. Le projet ReLIRE (Registre des livres indisponibles en réédition électronique) permet ainsi de numériser dans le respect des droits de propriété intellectuelle des ouvrages indisponibles parus avant 2001.

Le dépôt légal des supports nativement numériques

Les collections numériques de la BnF se doivent, en outre, de faire la jonction entre la numérisation rétrospective du patrimoine et l'avenir des supports de l'édition, qui seront de plus en plus numériques.

Au titre du dépôt légal, le département de l'audiovisuel reçoit et conserve des enregistrements sonores, des vidéos, des documents multisupports et des documents électroniques. Les premiers documents audiovisuels collectés sous forme numérique l'ont été dans les années 1980, à la faveur de la commercialisation des disquettes et cassettes logicielles, puis des supports optiques. En communiquant, dès 1996, des documents audiovisuels sur serveur vidéo numérique, la BnF a été pilote dans l'expérimentation de la communication numérique de documents. Avec l'apparition du numérique dématérialisé et l'évolution très rapide des chaînes d'édition et de diffusion, la Bibliothèque a adapté ses processus d'entrée. Elle réceptionne actuellement des fichiers dématérialisés, dans des proportions différentes selon les types de médias, mais qui peuvent atteindre jusqu'à 40 % des dépôts. Au printemps 2016, les documents audiovisuels numériques natifs issus du dépôt légal représentent ainsi plus de 25 000 documents. La volumétrie de cette collection est amenée à s'accroître de manière très conséquente dans les prochaines années à la faveur de la mise en place d'une filière automatisée d'entrée par flux des fichiers audiovisuels. La première filière développée concernera les documents sonores.

La BnF a pu constituer, grâce à l'archivage du web, une collection de plus de 560 téraoctets de données.

En 2006, la loi relative au droit d'auteur et aux droits voisins dans la société de l'information, dite loi DADVSI, a étendu le dépôt légal à tous les « signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique ». La BnF a ainsi pu constituer, grâce à l'archivage du web, une collection de plus de 668 téraoctets de données, qui permet, par exemple, aux chercheurs d'analyser l'utilisation d'internet comme support pour le développement de mouvements militants, ou de revivre la révolution tunisienne de 2010-2011 en s'intéressant à l'impact qu'elle a pu avoir dans les médias numériques et sur les réseaux sociaux.

En 2016, le champ du dépôt légal s'étend aux livres numériques, qu'ils soient homothétiques d'ouvrages imprimés déposés ou qu'ils n'existent que sous cette forme. Après une première expérimentation, une chaîne de traitement complète permettra de faire entrer plusieurs milliers de titres dans les collections dès la première année.

Les acquisitions électroniques

Ce nouveau patrimoine doit enfin être appréhendé dans le contexte plus large du développement des savoirs et des services numériques. Dès 1995, la BnF s'engage dans **une politique d'acquisition de trois types de ressources numériques natives :**

- **les bases de données :** ouvrages de référence (dictionnaires, encyclopédies, répertoires et annuaires), corpus de textes, portails spécialisés (164 abonnements ont été souscrits en 2015 ce qui représente 204 titres de bases de données proposées aux lecteurs) ;

- **les périodiques :** près de 1600 titres courants sont concernés. Accessibles en ligne, ils sont couplés ou non à un abonnement papier, les abonnements exclusivement électroniques relevant principalement des sciences et techniques ;

- **les livres numériques :** 91 900 titres ont été acquis auprès de grands éditeurs (Springer, Elsevier, Wiley, Royal Society of Chemistry), concomitamment à l'acquisition d'accès pérennes à des livres numériques vendus à l'unité (2 120 titres depuis 2012).

Devant l'accélération de la dématérialisation et des acquisitions numériques, la France s'est engagée dans une politique d'acquisitions numériques sous licence nationale, dont la BnF bénéficie pour les archives des revues. La Bibliothèque a, par ailleurs, intégré fin 2013 le Consortium universitaire de publications numériques (Couperin). Elle peut ainsi prendre part aux négociations nationales relatives à l'achat de nouvelles ressources et faire bénéficier ses lecteurs de l'accès à distance aux ressources acquises.

Une collection encyclopédique

La politique documentaire numérique de la BnF est complexe du fait des nombreux critères qui la déterminent : principes de sélection, contraintes techniques liées aux supports et à la nature des collections, sources de financement, sédimentation historique des politiques de numérisation, opportunités partenariales, émergence de nouvelles modalités de transmission des connaissances… Elle n'en est pas moins portée par un projet d'ensemble, inscrit dans un horizon d'exhaustivité « relative » ou « raisonnée » : l'ambition d'offrir au lecteur d'aujourd'hui, comme au chercheur de demain, une vision aussi complète que possible du patrimoine national. L'excellence de cette politique réside dans le fait de couvrir largement toutes les disciplines et de servir aussi bien les usages actuels que des besoins futurs qu'on ne peut anticiper. Elle s'opère sur un mode collaboratif, partenarial, avec une ambition nationale et internationale : enrichir toujours davantage les collections numériques pour les diffuser le plus largement possible.

L'ambition est d'offrir, au lecteur d'aujourd'hui comme au chercheur de demain, une vision aussi complète que possible du patrimoine national.

BnF

01.A Le dépôt légal du web

Institué par la loi du 1er août 2006, le dépôt légal du web tient une part importante dans l'accroissement des collections numériques de la BnF. L'ensemble des publications françaises en ligne est susceptible de faire l'objet d'une collecte par les établissements dépositaires (l'Ina - Institut national de l'audiovisuel - pour les sites de la radio et de la télévision, la BnF pour l'ensemble des autres sites français). La masse des documents disponibles sur internet et leur renouvellement interdisent l'exhaustivité ; la notion de représentativité régit la collecte, avec l'ambition d'offrir aux générations futures un miroir de la production culturelle française en ligne, au fil des époques.

Pour constituer des collections représentatives, la Bibliothèque a défini un modèle intégré qui mêle collectes larges (ensemble des sites français, une fois par an) et collectes ciblées de sites sélectionnés par des agents de l'établissement ou par des partenaires. Le dispositif a été enrichi, étendu et partagé (aux bibliothèques de dépôt légal imprimeur, notamment), afin de remplir plus efficacement ses objectifs et pour mieux servir les intérêts des utilisateurs des collections, chercheurs, amateurs, et toute personne intéressée.

LA QUESTION DE LA SÉLECTION

Le taux de couverture de la collecte large est encore insuffisant, même si l'adoption de méthodes de repérage automatique et la coopération avec d'autres acteurs nationaux de l'internet ont permis d'étendre cette couverture. La BnF coopère avec l'AFNIC (le registre qui gère le .fr) et complète la liste du .fr avec des sites automatiquement identifiés comme français (.com, .org, .net) : elle est ainsi en mesure de repérer et d'archiver 4 millions de sites.

Les collectes ciblées, touchant 20 000 sites à l'heure actuelle, peuvent être améliorées, avec une sélection plus largement ouverte aux partenaires de la BnF, à l'instar de la Bibliothèque nationale et universitaire de Strasbourg, qui a entrepris la collecte de son patrimoine local sur le web, les « Alsatiques » (voir axe 4 fiche I), via la BnF.
La réflexion sur la collecte des sites de l'État et des collectivités territoriales, lancée avec le service interministériel des Archives de France (SIAF), s'inscrit dans cette démarche. A d'ailleurs été mis en place un réseau national de sélection complétant le réseau de sélectionneurs du dépôt légal du web déjà présents au sein de la Bibliothèque.

LA QUESTION DE LA CONSERVATION

La conservation de ces contenus hétérogènes et complexes est primordiale. Les collections les plus récentes font l'objet d'un versement et d'une préservation dans SPAR, l'entrepôt numérique de la BnF. Il s'agit à présent d'y faire entrer les collections rétrospectives, ces véritables « incunables » du web de la fin des années 1990 et du début des années 2000, qui exigent des conditions de préservation optimales.

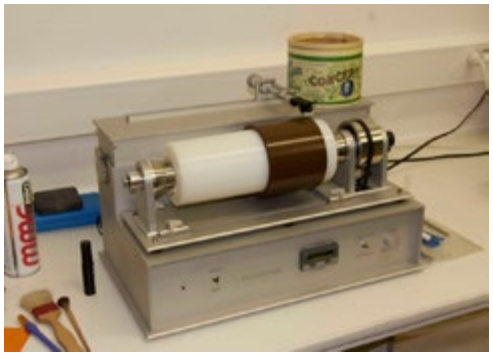
La notion de représentativité régit la collecte, avec l'ambition d'offrir aux générations futures un miroir de la production culturelle française en ligne, au fil des époques.

La presse quotidienne régionale (PQR)

**La PQR constitue une source essentielle de l'histoire locale, nationale et internationale. Parmi ces titres, 25 sont aujourd'hui moissonnés par le dépôt légal du web et correspondent à plus de 200 éditions locales. Sont ainsi concernés *les Dernières nouvelles d'Alsace* ou la *Dépêche du Midi*, ou encore le quotidien en ligne *Mediapart*.
Les journaux collectés par les robots du dépôt légal du web sont le plus souvent constitués de fichiers PDF, dont le contrôle de qualité et d'exhaustivité (bulletinage) est assuré quotidiennement par les mêmes agents que pour la presse imprimée.**

01.B La numérisation des documents audiovisuels

Riche de plus d'1,5 million de documents, la collection de documents audiovisuels de la BnF s'accroît chaque année d'environ 40 000 documents au titre du dépôt légal, mais aussi des dons et des acquisitions.
Afin d'assurer ses missions de conservation et de communication à long terme, la BnF a entrepris en 2000 un plan de numérisation dit « de sauvegarde » de ses collections patrimoniales sonores, audiovisuelles et multimédias, documents « en danger » à plusieurs titres : fragilité et autodégradation des supports, obsolescence matérielle et logicielle des équipements de lecture. Cette opération systématique et de grande ampleur concerne avant tout la numérisation des supports les plus fragiles : disques à gravure directe (« Pyrals »), cylindres de cire, supports magnétiques (bandes, cassettes, disquettes) et optiques (CD audio, DVD, cédéroms), pour en transférer les contenus dans des formats normalisés et pérennes. Les documents audio-analogiques sont numérisés en haute définition, sans compression et sans correction, de manière à conserver toute l'information contenue sur leurs supports. Les documents nés numériques sont recopiés bit par bit à l'identique. Le département de l'audiovisuel fait appel à des prestataires extérieurs pour la numérisation « de masse » de formats standards ou pour traiter des supports plus rares pour lesquels l'établissement ne conserve pas d'appareils de lecture (fils magnétiques, bandes Tefi, bandes vidéo 1 ou 2 pouces…).



Numérisation des cylindres du palais princier de Monaco avec l'Archéophone.

LES STUDIOS DE NUMÉRISATION AUDIOVISUELLE

La BnF dispose de quatre studios de numérisation pour le son (sites François-Mitterrand et Bussy) et de deux pour la vidéo (Bussy). Ils prennent en charge les documents ne pouvant faire l'objet de traitements de masse : disque à gravure directe, cylindres. Enfin, une filière interne permet la numérisation des supports DVD.

Ces studios assurent en partie la restauration du signal. Une filière de restauration de vidéogrammes a été mise en place en 2014 afin de traiter les documents en péril.

L'EXPERTISE AUDIOVISUELLE

Le département de l'audiovisuel met son expertise au service des collections de la BnF, de celles qui lui sont confiées en dépôt et parfois de collections extérieures traitées sur place dans le cadre de prestations de services (numérisation des cylindres du palais princier de Monaco, par exemple). Ainsi s'est affirmé au fil des ans un savoir-faire de référence en matière de numérisation et de restauration de vidéogrammes et de phonogrammes.

Un partenariat pour la numérisation des microsillons

Un projet conduit par BnF-Partenariats depuis 2012 a permis la numérisation de 46 000 disques microsillons (de 1949 à 1962 inclus), soit plus de 300 000 titres tous genres confondus (classique, jazz, variété, musiques du monde…).
Les microsillons ont été numérisés en 2014 par le partenaire technique Memnon Archiving services. Plus de 36 000 albums ont été mis en ligne par Believe Digital, après un travail d'enrichissement des métadonnées, sur les plateformes de diffusion de musique en ligne (téléchargement ou streaming, iTunes, Qobuz, Deezer, Spotify…).
Tous les contenus sont disponibles dans les salles de lecture de la BnF, et par extraits dans Gallica.
Les premiers revenus générés par la commercialisation des microsillons vont permettre la numérisation d'un corpus d'un millier de 78 tours, portant sur la chanson française dans le cinéma des années 30.

Au fil des ans, s'est affirmé un savoir-faire de référence en matière de numérisation et de restauration de vidéogrammes et de phonogrammes.

01.C La politique documentaire de la numérisation

La BnF s'attache à soutenir une orientation multisupport de la numérisation, reflétant la diversité du patrimoine dont elle a la garde : manuscrits, livres, périodiques ou presse, gravures, dessins, photographies, cartes, partitions, documents sonores, monnaies et images animées trouvent ainsi leur place dans Gallica.

La numérisation procède selon plusieurs logiques : une logique exhaustive lorsque l'on choisit de numériser tout un fonds ; une logique sélective lorsque sont établis des corpus thématiques, des ensembles intellectuels rassemblant des documents ayant un point commun ; une numérisation à la pièce, minoritaire, faisant droit aux demandes des publics de la BnF.

LA NUMÉRISATION DES LIVRES ET REVUES

Elle porte depuis 2007 sur les documents en langue française ou langues régionales qui ne sont plus protégés au titre de la propriété intellectuelle, et ont été imprimés sur le territoire français. La qualité du matériel de prise de vue n'ayant cessé de s'améliorer, les possibilités de sélection se sont élargies. Les campagnes de numérisation, s'étendent désormais aux ouvrages des XVIIe et XVIIIe siècles ainsi qu'aux ouvrages précieux ou fragiles. Parallèlement, la constitution de deux ensembles documentaires se poursuit : la numérisation des grands textes (sélection d'un ensemble de près de 10 000 éditions des œuvres emblématiques classées par thème et par siècle) et la numérisation des documents les plus consultés, représentant un ensemble de 5 000 documents par département thématique (Histoire philosophie sciences de l'homme, Droit économie politique, Sciences et techniques, Littérature et art). La numérisation de la presse reste enfin une priorité.

LES PROGRAMMES DE NUMÉRISATION DES COLLECTIONS SPÉCIALISÉES

S'ils mettent l'accent sur la reproduction des « trésors et chefs d'œuvre » (estampes de Pierre Roche, de Berthe Morisot ou encore de Vincent Van Gogh, documents du fonds Marcel Proust au département des Manuscrits, programme Rameau au département de la Musique ou encore numérisation des manuscrits enluminés médiévaux de la réserve de la Bibliothèque de l'Arsenal…), ces programmes s'étendent aussi à des ensembles de documents sélectionnés pour leur intérêt documentaire, fréquemment en lien avec des programmes de recherche. On peut citer, par exemple, le chantier de numérisation des monnaies qui a d'ores et déjà permis de mettre en ligne un quart des 700 000 monnaies conservées à la BnF. Des chantiers spécifiques sont montés en fonction des partenariats et accords passés par la Bibliothèque nationale de France dans des domaines très différents, comme la numérisation en 3D de 55 globes du département des Cartes et plans par la société DNP (Dai Nippon Printing) dans le cadre d'un mécénat de compétence (voir axe 6 fiche H).

Dans le cadre de son contrat de performance 2014-2016, la BnF a mis en chantier une charte documentaire générale de la numérisation.

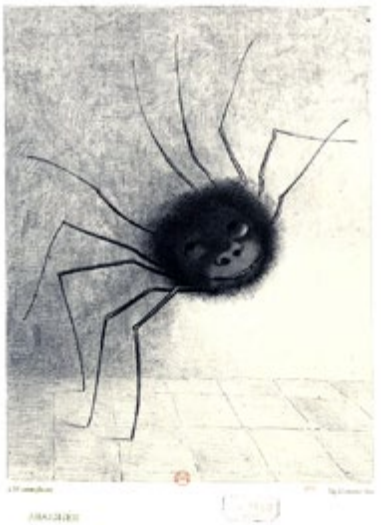
Dans le cadre de son contrat de performance 2014-2016, la BnF a mis en chantier une charte documentaire générale de la numérisation, afin de mieux faire connaître sa politique dans ce domaine, en lien avec les communautés d'usagers et les axes de coopération nationale et internationale.

Trois grands principes fondateurs

Les principes qui président à la décision de numériser un document sont de trois ordres :

- **l'intérêt patrimonial** : il s'applique à tous les documents auxquels la qualité esthétique, la valeur historique, la provenance, la rareté confèrent un statut de chef d'œuvre ou de trésor ;
- **l'intérêt documentaire** : sanctionné par les pratiques de recherche, il caractérise les ensembles très demandés ;
- **l'exigence de conservation** : elle motive la numérisation des documents fragilisés, ou dont l'état rend la communication difficile ou risquée : documents acidifiés (en particulier la presse), grands formats, plaques de verre, objets avec des éléments mobiles… En matière de conservation, l'audiovisuel est un cas particulier, les supports magnétiques, optiques et mécaniques étant fragiles et menacés de dégradation accélérée.

Araignée (estampe), par Odilon Redon



01.D Les différents modes de financement de la numérisation

L'ensemble des possibilités de financement ont été mises en œuvre. Ces dispositifs se cumulent pour donner à la BnF une capacité d'action d'une ampleur inégalée dans le domaine de la numérisation du patrimoine – même si, ponctuellement, sur certains fonds comme la presse, des besoins restent à satisfaire.

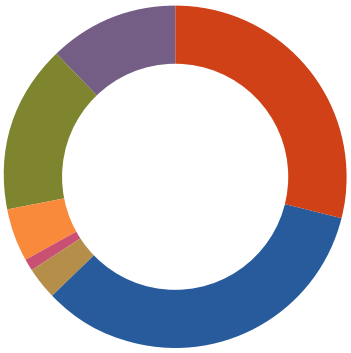
FINANCEMENT PAR FONDS PROPRES

(à partir de 1997)
17 millions d'euros de dotation budgétaire ont été affectés par la BnF de 2002 à 2012 à la numérisation, dont 3,2 vers les partenaires et pôles associés.

FINANCEMENT PAR CRÉDITS DU CENTRE NATIONAL DU LIVRE (CNL)

(à partir de 2007, pour les imprimés)
46,7 millions d'euros ont été reçus de 2007 à 2014 (moyenne annuelle proche de 6 millions d'euros)

RÉPARTITION DES VUES NUMÉRISÉES ET POST-PRODUITES - 2015



pour financer à la fois la numérisation des imprimés, la conservation numérique dans SPAR, la diffusion dans Gallica et, depuis peu, le projet ReLIRE.

FINANCEMENT PAR ACTIONS DE MÉCÉNAT D'ENTREPRISE

Plusieurs projets ont bénéficié d'un apport depuis 2009 : par exemple 150 000 euros pour la numérisation de 350 manuscrits arabes, turcs et persans soutenue par la Fondation Total en 2009, ou la numérisation et la restitution en 3D, en très haute définition, d'un ensemble de globes terrestres et célestes du XIe au XIXe siècle, soutenue par un mécénat de compétence de Dai Nippon Printing en 2015 (voir axe 6 fiche H).

FINANCEMENT PAR ACTIONS DE MÉCÉNAT INDIVIDUEL

L'opération « Adoptez un livre » a été initiée et lancée par l'Association des Amis de la BnF en 2011. À fin 2015, plus de 300 ouvrages avaient été numérisés par ce biais.

FINANCEMENT PAR PROJETS

Depuis 2009, des projets collaboratifs ont vu le jour au niveau national et international, financés par des fonds privés, européens, ou liés au Grand Emprunt. Ces projets sont l'occasion d'associer les bibliothèques françaises à la constitution d'ensembles documentaires, par la reconstitution de bibliothèques anciennes aujourd'hui dispersées (l'International Dunhuang Project, Le Roman de la Rose, BHL-Europe, Europeana Regia, le LABEX OBVIL et le programme de numérisation de manuscrits de Christine de Pisan).

FINANCEMENT VIA LA MOBILISATION DES « INVESTISSEMENTS D'AVENIR »

Afin de pouvoir émarger à ces crédits du Commissariat général à l'investissement (CGI), la bibliothèque a créé une filiale, BnF-Partenariats, qui pilote désormais des partenariats public-privé sur des projets de numérisation de fonds spécifiques.

Par ailleurs, dans le cadre de la numérisation des livres indisponibles du XXe siècle, le montage prévoit un triple financement CNL, CGI et éditeurs (notamment le Cercle de la Librairie).

PRESTATIONS DE REPRODUCTION

Elles concourent à la numérisation des fonds : un million de vues sont réalisées chaque année dans le cadre de 13 000 commandes de particuliers et de professionnels. Pour l'année 2015, les recettes liées aux travaux de reproduction se sont élevées à 1 104 179 euros pour les travaux, 541 683 euros pour la redevance d'utilisation commerciale et 21 402 euros pour les partenariats, soit un total de 1 667 265 euros TTC.

BnF-Partenariats

Cette filiale créée en 2012 a permis d'initier plusieurs accords de partenariat pour la numérisation :

- **avec ProQuest, en vue de la numérisation d'environ 28 000 livres anciens imprimés entre 1470 et 1700 (soit 10 millions de pages) ;**
- **avec Believe Digital et Memnon Archiving Services, en ce qui concerne la numérisation et la restauration du signal sonore de près de 46 000 microsillons ;**
- **avec Immanens, pour accélérer la numérisation et la diffusion en ligne des fonds de presse écrite, depuis les origines de celle-ci.**

01.E
Les ateliers internes de numérisation

Trois ateliers internes de reproduction, situés sur les sites de Bussy-Saint-Georges, Sablé-sur-Sarthe et Paris (site François-Mitterrand) sont dédiés à la reproduction de conservation. Amorcé en 2005, leur basculement progressif de l'argentique (microfilm et microfiche) vers le numérique est aujourd'hui quasiment achevé.
Un parc de plus de 25 scanners permet de traiter une grande diversité de documents (textes, éléments graphiques, objets), issus des différents départements de la direction des collections et ne pouvant être couverts par les marchés de numérisation : les grands formats (à Sablé), les documents très fragiles et précieux (Paris et Sablé), la presse fragile (Bussy), les objets (atelier mobile) et la procédure dite « urgent lecteur numérique » sur le site de Paris.

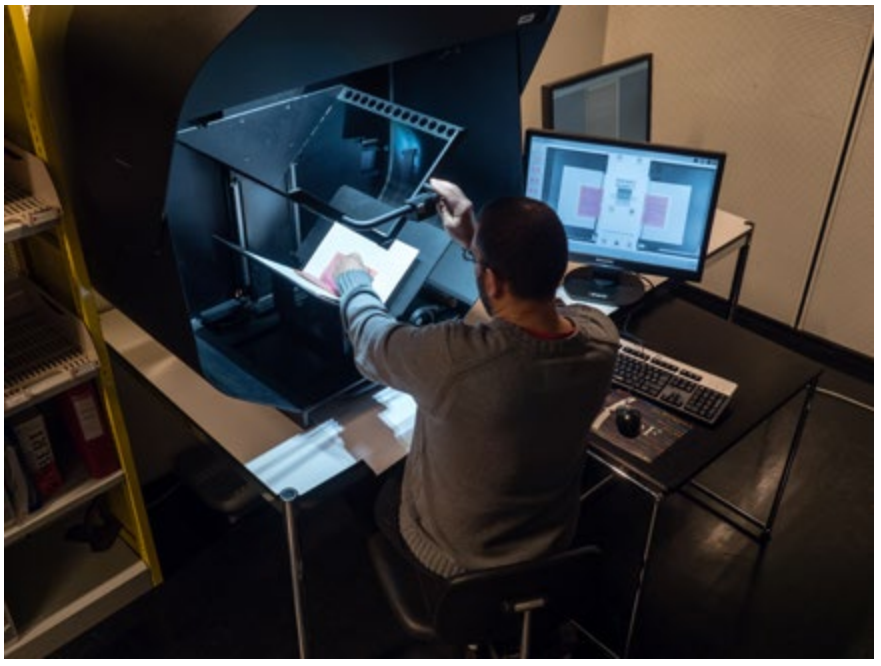
À ces ateliers s'ajoutent quatre studios de numérisation pour le son (sites François-Mitterrand et Bussy) et deux studios de numérisation pour la vidéo (Bussy).

Chaque site a ainsi développé une spécialisation technique – ou un domaine d'excellence –, sans préjudice d'autres compétences. La qualité du travail fourni par les ateliers tient à la mise en place de synergies avec les chargés de collections, les ateliers de restauration et ceux de préparation physique, voire avec les catalogueurs. Soulignons aussi la polyvalence du parc de matériels, qui offre de nombreuses solutions de traitement.

Sur les sites François-Mitterrand et Richelieu, deux ateliers de numérisation et d'impression à la demande sont chargés de traiter chaque année environ 13 000 commandes de clients souhaitant acquérir une reproduction dans le cadre d'un projet professionnel, personnel ou de recherche (voir axe 3 fiche E).

Chaque site a développé une spécialisation technique ou un domaine d'excellence.

Recours à des machines qui permettent la numérisation en ouverture partielle : ici, le Qidenus



01.F
Le cas particulier de la presse

Source essentielle pour la recherche, l'enseignement pédagogique et l'information, la presse, destinée à l'origine à un usage éphémère, pose des problèmes spécifiques de conservation et de sauvegarde. Le papier et l'encre, de mauvaise qualité, accélèrent la détérioration du document : la pâte à bois, matériau largement employé, rend le papier très acide. La reproduction des fascicules de presse est donc cruciale à double titre.

Les campagnes de numérisation successives ont rendu partiellement accessible, dans Gallica, un corpus de plusieurs dizaines de titres : grands quotidiens nationaux, presse spécialisée, presse locale. Avec plus de 5 millions de pages de presse numérisées depuis 2004 (dont près de 4 millions océrisées), la BnF propose sur le web l'une des plus importantes hémérothèques numériques en Europe.

UN GISEMENT INEXPLORÉ

En raison de l'essor considérable de la presse en France au milieu du XIXe siècle, les besoins excèdent les moyens consacrés aux opérations de reproduction : seuls 10 % des titres sont à ce jour microfilmés, et moins de 3 % numérisés, tandis que seules 2 500 des 110 000 publications conservées par l'institution, représentant a minima 100 millions de pages, sont, à l'heure actuelle, disponibles en ligne.

DES CHOIX DOCUMENTAIRES FONDÉS SUR LES USAGES

Les choix documentaires prennent en compte les statistiques de consultation des journaux sur place et les suggestions des utilisateurs individuels comme des écoles doctorales.

L'accent est également mis sur les collections appartenant au domaine public (imprimés avant 1945), en vue d'un affichage immédiat dans Gallica. Une ouverture à la période postérieure n'en est pas moins à encourager, en liaison avec les éditeurs et les ayants droit.

Les nouvelles techniques d'enrichissement des fichiers numériques (segmentation à l'article, entités nommées, rubriquage, correction des titres) font l'objet d'un second niveau de sélection, pour retenir le ou les types de presse qui doivent prioritairement bénéficier de ces améliorations.

La BnF propose sur le web l'une des plus importantes hémérothèques numériques en Europe.

La presse régionale

En 2013, la BnF a mené une enquête sur la numérisation de la presse en région. Près d'un tiers des bibliothèques municipales classées et services d'archives départementales ont indiqué avoir numérisé près de 5,5 millions de pages de presse, soit environ un millier de titres. Le financement provient en grande partie du Plan national de numérisation, parfois avec un appoint des pôles associés régionaux de la BnF (indexation de la presse numérisée en Languedoc-Roussillon).
Le site presselocaleancienne.bnf.fr, lancé en 2016, constitue une base des données enrichie de liens numériques vers les titres de la presse ancienne, locale et régionale, numérisée en France.

01.G La numérisation des œuvres encore soumises à la propriété intellectuelle

Dans le cadre de la coopération numérique nationale, de nombreux documents encore protégés par le droit d'auteur sont diffusés dans Gallica après négociation de droits et à la faveur d'accords spécifiques de numérisation et de diffusion.

QUELQUES EXEMPLES

C'est le cas de plusieurs revues contemporaines importantes dans le domaine des sciences humaines et sociales (*Études, Revue française de psychanalyse, Revue historique, Revue d'histoire littéraire de la France, Revue d'histoire moderne et contemporaine, Le mouvement social, Le Moyen Âge, Hérodote, XVIIIe siècle*). C'est aussi le cas de publications de sociétés savantes locales et régionales (100 conventions de cession de droits), d'académies, de sociétés nationales ou régionales (sciences, médecine, agriculture, etc.). Gallica donne également accès aux publications officielles à partir des collections de la BnF (Journal officiel, procès-verbaux de conseils généraux, etc.) ou de celles de ses partenaires (débats parlementaires de l'Assemblée nationale et du Sénat, INSEE, Banque de France, Observatoire de Paris, Ville de Paris, Cedias-Musée social, Institut catholique de Paris, etc.).

RENDRE ACCESSIBLES DES OUVRAGES INCOMMUNICABLES

La numérisation d'ouvrages sous droits permet également de rendre à nouveau accessibles, via Gallica intra muros, consultable dans les salles de lecture de la bibliothèque de recherche, des ouvrages devenus incommunicables pour des raisons de conservation.

La sélection des ouvrages à numériser s'inscrit en règle générale dans un plan de conservation reposant sur des vérifications systématiques de l'état de conservation des ouvrages en magasin. Une procédure de numérisation spécifique peut également être lancée dans le cadre du circuit « urgent lecteur », qui permet de communiquer sous la forme d'un support de substitution un document incommunicable sur son support d'origine, dans un délai maximal d'un mois entre la demande du lecteur et la mise en ligne sur Gallica ou Gallica intra muros. 130 000 œuvres encore protégées au titre de la propriété intellectuelle ont ainsi été numérisées par l'établissement.

La sélection des ouvrages à numériser s'inscrit dans un plan de conservation.

ReLIRE

Le projet ReLIRE est consécutif à la loi du 1er mars 2012 relative à l'exploitation numérique des livres indisponibles du XXe siècle. Il vise à encourager la réédition et la commercialisation en format numérique de livres non encore entrés dans le domaine public.
La sélection des livres appelés à figurer dans le registre ReLIRE est assurée par un comité scientifique où représentants des auteurs et des éditeurs siègent à parité. Publiée annuellement par la BnF, la liste des ouvrages sélectionnés est consignée dans un registre librement accessible en ligne (relire.bnf.fr) qui comprend plus de 215 000 oeuvres au 21 mars 2016 et devrait identifier 500 000 livres à l'horizon 2023.
La numérisation des livres, issus des collections de la BnF, a débuté en 2014 avec 15 000 ouvrages. Une société de projet, FeniXX, filiale du Cercle de la Librairie créée à cet effet, assure la diffusion commerciale des livres numériques.
Les ouvrages numérisés sont rendus accessibles parallèlement dans Gallica intra muros et, dès leur mise en vente, dans Gallica, à raison d'un extrait des 15 premiers pour-cent.
À leur entrée dans le domaine public, les livres du registre ReLIRE seront rendus accessibles dans Gallica intégralement et gratuitement.

01.H
Le dépôt légal des livres numériques

Les livres nativement numériques sont proposés et vendus en ligne : ils entreront à la BnF par dépôt légal. Dès 2016, les premiers fichiers numériques intègrent les collections à titre expérimental.

UN CIRCUIT DE DISTRIBUTION SPÉCIFIQUE

À la différence de celle des sites web, la collecte des livres numériques ne peut pas être effectuée automatiquement en ligne. Le circuit de distribution de ces ouvrages conduit en effet à ce que le propriétaire du site web par lequel la vente se fait (celui du libraire en général) ne soit jamais en possession du fichier du livre, qui reste entre les mains du distributeur. Il a donc été décidé de créer un mode d'entrée spécifique, se rapprochant du modèle du dépôt des ouvrages imprimés.

UN CIRCUIT DE DÉPÔT LÉGAL ADAPTÉ

En écho à ce circuit de distribution, un circuit complet de dépôt légal a été conçu. Si les éditeurs gardent, devant la loi, la responsabilité du dépôt légal, la possibilité leur est donnée de mandater leur distributeur en ligne aux fins d'effectuer le dépôt en leur nom. Cette collaboration facilite la levée des mesures techniques de protection (DRM) apposées par les distributeurs, qui empêchent la gestion et la bonne conservation des livres par la BnF. Il en est de même des livres vendus dans un format propriétaire, que la BnF ne peut pas réutiliser et préserver librement : ce sont en effet les distributeurs qui disposent des ouvrages aux formats ouverts PDF et ePub, avant qu'ils ne soient convertis dans des formats propriétaires.

En outre, afin d'être exploitables par les plateformes de vente en ligne, les livres numériques sont décrits dans des fichiers de métadonnées (fichiers ONIX) mis en forme par les distributeurs à partir des informations fournies par les éditeurs : la récupération de ces fichiers vient enrichir la chaîne de traitement des livres numériques par la BnF.

La chaîne de traitement

Le circuit de traitement des livres numériques entrant par dépôt légal à la BnF a été pensé en fonction des applications et des outils existants. L'extranet du dépôt légal assure le traitement des métadonnées qui seront ensuite versées dans le catalogue général.
Le système de préservation numérique SPAR (voir axe 2 fiche B) gère la conservation à long terme des fichiers. Seule la chaîne d'entrée des documents numériques (CEDN) doit être intégralement créée, le reste des outils faisant l'objet d'adaptations afin de pouvoir accueillir ces nouveaux documents. L'intégralité du processus est conçue pour fonctionner de manière automatisée.
Afin de sécuriser l'accès aux collections, la consultation se fera via Gallica intra muros dans les salles de recherche.

L'intégralité du processus de traitement est conçue pour fonctionner de manière automatisée.

BnF

02

SIGNALER, CONSERVER

13 millions
de notices bibliographiques et 2,5 millions de notices d'autorité dans les catalogues

2,4 millions
de pages dans data.bnf.fr dont 2 millions de pages auteur

8,5 millions
de documents liés aux pages de data.bnf.fr

813 000
documents audiovisuels dans le système audiovisuel (soit 1,7 pétaoctets)

32 Po
Capacité théorique des librairies bandes de SPAR (2 sites)

L'expertise séculaire développée par les bibliothèques dans le traitement des collections s'applique particulièrement dans deux domaines, vécus comme le cœur de métier des bibliothécaires : le signalement, c'est-à-dire l'ensemble des techniques permettant d'identifier, de décrire et de trouver un document, et la conservation. À l'heure du numérique, la BnF doit, pour mener à bien les missions de conservation et de diffusion du patrimoine dont elle a la charge, développer et parfois réinventer cette expertise en l'appliquant à de nouveaux objets, de nouveaux supports, de nouveaux usages. Les catalogues de la BnF, riches de plus de 13 millions de notices, ont entamé leur révolution vers le web des données ouvertes. L'entrepôt numérique SPAR (Système de préservation et d'archivage réparti), qui contient aujourd'hui plus de 5 millions de documents, s'est ouvert au tiers archivage et a obtenu en 2013 l'agrément du SIAF (Service interministériel des Archives de France).

ZOOM SUR QUELQUES POINTS FORTS

02.A_ Data.bnf.fr
02.B_ L'entrepôt d'archivage numérique SPAR

MOYENS, OUTILS ET MÉTHODES

02.C_ L'évolution du catalogue général vers le modèle FRBR
02.D_ Ressources électroniques acquises : signalement et accès
02.E_ Garantir une consultation pérenne des documents audiovisuels

AXES DE DÉVELOPPEMENT

02.F_ La transition bibliographique

L'évolution du signalement à l'heure des données ouvertes

Les données de description bibliographique ou « métadonnées » sont au cœur du métier de bibliothécaire. Elles irriguent tous les processus, physiques ou numériques, qui permettent à une bibliothèque de conduire ses missions, du magasinage à la mise en valeur de ses collections en passant par les traitements de conservation ou de numérisation. C'est la raison pour laquelle la BnF doit adapter en profondeur ses outils et ses processus de signalement si elle veut accompagner efficacement le développement des activités numériques.

À l'heure des mégadonnées (*big data*), les métadonnées qualifiées, fiables et structurées constituent un atout stratégique : elles permettent d'être repéré, référencé et finalement valorisé au milieu de la masse. La modernisation des dispositifs de production et de diffusion des métadonnées de la BnF devient ainsi le corollaire indispensable d'une politique numérique ambitieuse. Ce domaine est justement l'un des points forts de la BnF, qui a démontré sa capacité à créer et diffuser une masse critique de métadonnées de référence en s'appuyant sur ses acquis : un solide capital de confiance auprès de ses usagers et partenaires ; un trésor de millions de notices rigoureusement établies ; une expertise multisupport, adaptée à la richesse et à la diversité de la production culturelle nationale et internationale. Le mouvement d'ouverture des données publiques (*open data*) et le développement du web sémantique constituent à cet égard une réelle opportunité technologique pour le développement des missions de service public qui incombent à la BnF.

Le mouvement d'ouverture des données publiques (*open data*) et le développement du web sémantique constituent une réelle opportunité technologique pour le développement des missions de service public qui incombent à la BnF.

À travers des adaptations à la fois techniques (modèles et formats de métadonnées), informatiques (processus de traitement de données, applications de production et interfaces de diffusion) et organisationnelles (évolution des métiers et des activités), la BnF opère une véritable « transition bibliographique », caractérisée notamment par l'évolution progressive des catalogues traditionnels vers les moteurs de recherche du web, et par une nouvelle manière d'organiser l'accès à l'information.

La BnF entend également profiter de cette dynamique pour susciter et faciliter la réutilisation par des tiers de ses métadonnées, placées depuis janvier 2014 sous licence ouverte de l'État. Et l'on constate, de fait, une augmentation de la visibilité des ressources bibliographiques de la BnF dans les grandes bases de connaissances commerciales internationales, lesquelles alimentent nombre de systèmes de bibliothèques (OCLC – Ex Libris – EBSCO, etc.). Son ambition à long terme est de participer à un véritable service public des métadonnées culturelles, coconstruit avec des acteurs publics (Agence bibliographique de l'enseignement supérieur, Archives nationales…) et privés (éditeurs, start-ups, industries culturelles…), et proposant des services innovants à forte valeur ajoutée.

La conservation numérique

La conservation est l'une des principales missions de la BnF et s'inscrit, pour le numérique, dans la continuité de la conservation physique. La numérisation, avec l'abandon progressif des supports argentiques, devient l'unique modalité de reproduction des documents à des fins de conservation.

À partir des années 1980, la Bibliothèque nationale avait lancé de grandes campagnes de reproduction sur microformes de ses collections les plus fragiles, dans le cadre de plans de sauvegarde d'envergure. La numérisation, qui a pris le relais de la microforme, garantit une meilleure qualité d'image, une consultation plus aisée des documents reproduits, et ouvre la possibilité d'une consultation à distance. Cependant, la question de sa fiabilité en matière de conservation se pose encore : la pérennité des supports numériques est liée à leur soutenabilité, que ce soit sur le plan technique ou au regard des financements importants qu'ils mobilisent.

Par ailleurs, l'ampleur des moyens dévolus ces dernières années à la numérisation ainsi que l'enrichissement des collections par des masses de plus en plus importantes et hétérogènes de documents nativement numériques (voir axe 1) ont élargi la problématique de la conservation numérique. L'enjeu n'est plus seulement de sauvegarder des originaux menacés de dégradation ou de disparition en mettant un support de substitution à disposition, mais de prendre en compte le numérique natif comme relevant des collections en tant que telles, au même titre que les différents supports – physiques (manuscrits, imprimés, objets, disques microsillons), magnétiques (disquettes, cassettes vidéo ou audio), optiques (cédéroms, DVD) – qui constituent le patrimoine de la Bibliothèque dans toute sa diversité.

L'enjeu n'est plus seulement de sauvegarder des originaux menacés de dégradation ou de disparition, mais de prendre en compte le numérique natif comme relevant des collections en tant que telles.

Face à ce nouveau défi, la BnF a progressivement mis en place à partir de 2005 un système de pérennisation des données numériques (SPAR) qui n'a pas d'équivalent en France en ce qu'il promet, pour des volumes de données très importants (capacité actuelle de deux fois 16 pétaoctets), une longévité qui ne serait limitée ni par la durée de vie des supports, ni par celle de l'environnement technologique (logiciels, formats, systèmes d'exploitation, terminaux…). SPAR a été déployé en plusieurs étapes et continue à évoluer pour répondre aux besoins croissants de sauvegarde numérique. C'est un système ouvert et modulaire reposant sur la norme internationale OAIS (*Open Archival Information System*), capable notamment d'évoluer au rythme de l'obsolescence des environnements informatiques.

Les solutions techniques ne permettent cependant pas seules d'améliorer la gestion des collections numériques et d'assurer leur pérennité. L'affectation des moyens aux différentes filières d'entrée dans les magasins numériques de SPAR repose sur des critères scientifiques : priorité aux documents les plus fragiles, gestion des risques différenciée en fonction de la vulnérabilité des documents, de leur rareté ou de l'importance de leur contenu… La pérennisation de la consultation, afin de garantir l'authenticité des documents et la confiance dans le système de conservation, passe aussi par un important travail de documentation et d'archivage des environnements de lecture : lorsqu'on archive un jeu vidéo ou une page web, on ne peut se contenter de garder la mémoire de leur seul contenu. Conserver le patrimoine numérique, c'est aussi préserver l'épaisseur historique de son immatérialité : l'expérience sensorielle qui l'accompagne, le souvenir de la manière dont la culture et la connaissance imprègnent le quotidien d'une société.

02.A Data.bnf.fr

L'application pionnière data.bnf.fr concentre aujourd'hui les plus grands enjeux et suscite le plus d'attentes en matière d'ouverture des données. Si la BnF continuera d'avoir besoin d'applications distinctes, aux fonctionnalités adaptées à ses différents types de ressources (Gallica, le catalogue général, BnF archives et manuscrits…), data.bnf.fr vise à offrir un point d'entrée unique, depuis le web, à l'ensemble de ces ressources.

UN OUTIL INTÉGRÉ À L'ÉCOSYSTÈME DU WEB

Cet outil propose aux internautes des pages riches et organisées sur les sujets qui les intéressent (les auteurs, les œuvres, les lieux…), ainsi que de nombreux liens, aussi bien entre ces pages que vers d'autres sites. Cette logique, en prenant ses distances par rapport à l'organisation classique de l'information dans un catalogue de bibliothèque, favorise le référencement, génère du trafic pour les autres applications de la BnF et permet enfin de toucher un public considérablement plus large que celui qui fréquente déjà les bibliothèques.

UN SUCCÈS D'AUDIENCE

Lancé en 2011, data.bnf.fr connaît déjà un beau succès d'audience : sa consultation quotidienne a doublé entre fin 2014 et fin 2015 pour atteindre près de 10 000 visites quotidiennes, plaçant data.bnf.fr en quatrième position des sites de la BnF les plus consultés.

data.bnf.fr concentre aujourd'hui les plus grands enjeux et suscite le plus d'attentes en matière d'ouverture des données.

LA CLÉ DE VOÛTE D'UN NOUVEAU MODE DE COOPÉRATION

Reposant sur des technologies et des formats plus ouverts que ceux utilisés traditionnellement par les bibliothèques, il permet d'imaginer à terme des scénarios de coproduction de données avec des communautés professionnelles beaucoup plus diversifiées. Dans les années à venir, la BnF entend fédérer autour de data.bnf.fr un écosystème de partenaires producteurs et réutilisateurs de données culturelles, scientifiques ou éducatives, en particulier grâce à une politique active de liens et d'intégration de données issues de partenaires du domaine culturel. En 2016, data.bnf.fr aura atteint une maturité technologique, une complétude et une robustesse qui en feront la clé de voûte de nouveaux modes de coopération, de nouveaux usages et de nouveaux services. En 2017, sa mise à jour deviendra mensuelle.

Une démarche d'ouverture des données

La BnF a adopté en 2014 la licence ouverte de l'État préconisée par la mission Etalab, pour data.bnf.fr et l'ensemble des métadonnées descriptives qu'elle diffuse.

En favorisant des logiques de large diffusion et de libre réutilisation, y compris commerciale, de ses métadonnées, la BnF souhaite démontrer sa capacité à servir davantage d'usagers et à contribuer à l'invention de nouveaux services, qu'on pense aux applications en mobilité ou, demain, aux objets connectés.

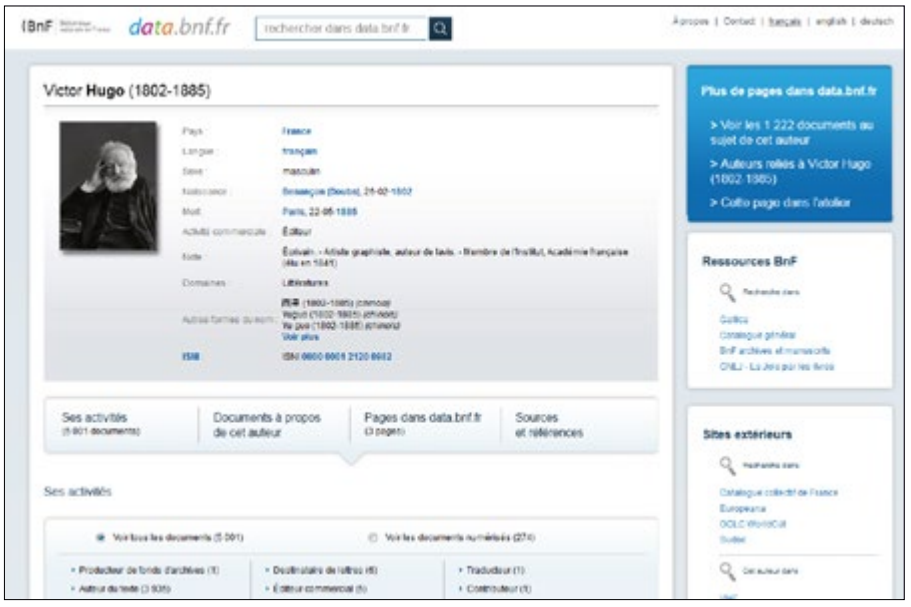
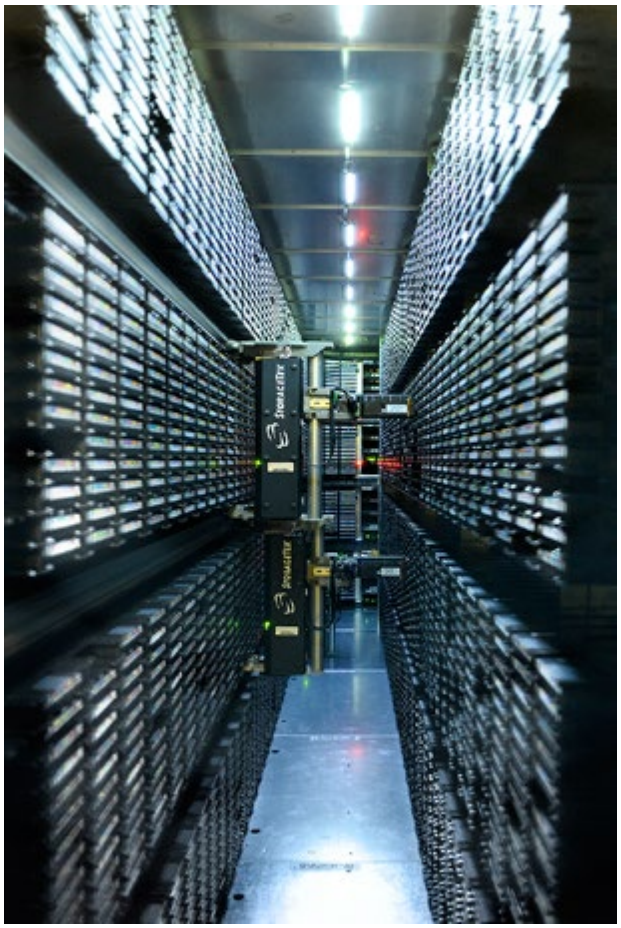


Schéma numérique - mars 2016

02.B L'entrepôt de préservation numérique SPAR

L'accroissement et la diversification des collections numériques de la BnF – plusieurs milliers de formats – posent de manière cruciale la question de l'entrepôt des objets numériques. La Bibliothèque a lancé le projet SPAR (Système de préservation d'archivage réparti). Véritable magasin numérique, SPAR est une réponse appropriée aux problématiques de préservation pérenne et sécurisée des données numériques. Initié en 2005, SPAR s'appuie sur des normes internationales faisant autorité : OAIS (*Open Archival Information System*) et PAIMAS (*Producer-Archive Interface Methodology Abstract Standard*), METS (*Metadata Encoding & Transmission Standard*), PREMIS (*Preservation Metadata Implementation Strategies*).



L'infrastructure du système SPAR : robotique de stockage sur bandes magnétiques

De mai 2010 à janvier 2016, 2,23 pétaoctets de données représentant 5 millions de paquets d'archives ont été versés dans SPAR.

PRÉVENTION ET GARANTIES

Non réductible à un simple entrepôt de données sécurisé qui ferait de multiples copies – trois copies sur deux sites – et assurerait une surveillance continue de l'état des supports d'enregistrement, SPAR permet, grâce à une reconnaissance précise et complète des formats de données versées, de garantir la continuité d'accès en procédant aux transformations nécessaires en cas d'obsolescence technologique des outils informatiques de restitution (logiciels, systèmes d'exploitations, matériels de lecture…). Le système de SPAR est organisé en filières (numérisation, dépôt légal du web, tiers archivage…), tenant compte de la diversité des collections qui peuvent y être conservées. De mai 2010 à janvier 2016, 2,23 pétaoctets de données représentant 5 millions de paquets d'archives ont été versés dans SPAR. Le 29 mars 2013, SPAR a reçu pour trois ans l'agrément du service interministériel des Archives de France pour la conservation d'archives publiques courantes et intermédiaires (publié au Journal officiel du 2 mai 2013).

SERVIR UNE COMMUNAUTÉ

SPAR n'a pas vocation à n'être qu'un outil interne : il a aussi été pensé pour être au service d'une communauté. Ainsi, la BnF a mis en place une solution de tiers archivage qui offre aux organisations les mêmes garanties de sécurité et de pérennité que celles dont bénéficient ses propres fichiers déposés dans SPAR (voir axe 4 fiche C). L'investissement consenti pour le développement du magasin numérique est important, que les coûts soient induits par l'accroissement des collections ou par la maintenance du système : la partie matérielle (disques et bandes) représente 500 000 à 1 million d'euros par an tandis que la partie logicielle représente 600 000 euros par an. Ces coûts sont supportés par le budget propre de la BnF et par des crédits du CNL. La recherche d'un troisième site d'ici 2020 permettrait de répondre à des risques de cataclysmes majeurs qui pourraient affecter simultanément les deux sites actuels, François-Mitterrand et Bussy-Saint-Georges (tempête solaire magnétique, accident nucléaire ou événement climatique majeur).

02.C L'évolution des catalogues : vers le modèle FRBR

La BnF a entrepris en 2013 une refonte de l'interface publique de son catalogue général visant à l'adapter aux usages du web et à proposer des points d'entrée spécifiques dans les collections (univers Jeunesse, Images et cartes, Musique). En 2014, s'est engagé en parallèle un chantier de refonte de l'interface publique du catalogue BnF archives et manuscrits au profit d'un dispositif beaucoup plus ergonomique et plus performant, dont l'ouverture est programmée en 2016.

UN MODÈLE ADAPTÉ AUX USAGES

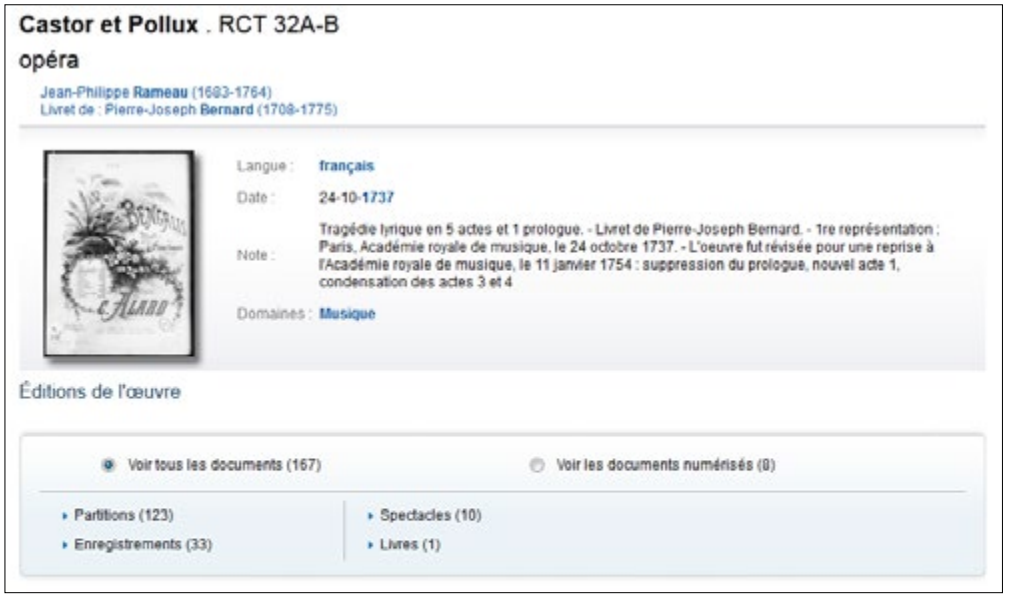
La structuration des données du catalogue est appelée à évoluer dans l'esprit de la recherche d'information sur le web, où l'internaute s'intéresse à des personnes, des concepts, des œuvres, des lieux, plus qu'il ne souhaite consulter des notices ou localiser des documents. Le modèle FRBR est la clé de voûte de cette réforme.

Sa finalité est double : présenter l'information bibliographique d'une manière plus intuitive pour l'usager et faciliter la gestion de l'information par le regroupement des différents types de publications associés à une même œuvre (traductions, livres numériques, dans différents formats, différentes éditions…).

UN CHANTIER À PART ENTIÈRE

Pour une institution comme la BnF, ce changement fondamental dans la structure des données implique d'abord un vaste chantier de traitement rétrospectif de son capital bibliographique, qui représente plus de 13 millions de notices. Ce chantier s'est engagé grâce aux technologies de data.bnf.fr, qui permettent d'ores et déjà, par l'exploitation d'algorithmes d'alignement et de regroupement, d'opérer un premier niveau de transformation des données.

Si ces traitements sont encore de nature expérimentale et nécessitent d'acquérir la maîtrise d'algorithmes complexes et l'appui d'experts afin de définir des seuils de confiance, leur reversement dans les catalogues sources est bien l'objectif final d'une telle démarche, qui permet de préparer la transition bibliographique.



Exemple de structuration de l'information suivant le modèle FRBR dans data.bnf.fr

Grâce au regroupement des différents types de publications associés à une même œuvre, l'information bibliographique est présentée d'une manière plus intuitive.

150 000 notices d'imprimés, de partitions, de documents sonores, de films et de documents électroniques et multisupports sont ainsi liées automatiquement à l'occasion du premier chargement dans le catalogue général début 2016. Les prochaines étapes concerneront notamment la création de notices d'oeuvres.

Le modèle FRBR

Il a été imaginé au sein des instances de normalisation internationale (IFLA) il y a près de vingt ans. En multipliant les liens entre différentes entités (auteurs, œuvres, sujets), il permet de préparer les données des bibliothèques à leur intégration dans le web sémantique. La FRBRisation des catalogues de bibliothèque s'effectue aujourd'hui le plus souvent de manière automatique, en regroupant les documents suivant des règles précises. L'évolution des normes de catalogage dans les années à venir permettra une FRBRisation plus systématique reposant sur des critères plus fiables.

02.D
Ressources électroniques acquises : signalement et accès

DEUX MODES D'ENTRÉE

Les ressources numériques acquises résultent de deux grands types d'opérations. Le premier est l'abonnement à des plateformes d'éditeurs, avec ou sans acquisition de fichiers ; le second consiste en des dons et acquisitions de fichiers numériques, pour lesquels l'accès au contenu doit être assuré par la BnF.

L'acquisition de documentation électronique auprès d'éditeurs peut s'effectuer à l'unité ou par bouquet. Elle concerne une grande diversité de contenus : livres numériques, périodiques, bases de données, mais aussi images, cartes, musique notée.
Ces ressources accessibles sur place et à distance selon les licences sont proposées aux lecteurs suivant deux protocoles : par liste alphabétique ou thématique (plus de 240 000 entrées étaient comptabilisées en décembre 2015 pour les seuls livres et les périodiques) ; par le moteur de recherche intégré d'EBSCO (EBSCO *Discovery Service*), qui ne permet toutefois d'interroger qu'une partie des ressources électroniques acquises par la Bibliothèque.

Les fichiers numériques qui ne sont pas accessibles sur les plateformes des éditeurs seront mis à disposition sur Gallica intra muros via la filière ADDN (Acquisitions et dons de documents numériques – voir axe 5 fiche I), en cours de développement.

UN ENJEU : LE SIGNALEMENT

Aujourd'hui, le signalement de ces ressources, sommaire et déconnecté du catalogue général, est difficilement visible pour les lecteurs. Leur taux de consultation demeure donc très faible, alors même que le quart du budget annuel des acquisitions courantes de la direction des collections est dédié à la seule documentation électronique, et qu'une partie des périodiques étrangers n'est désormais accessible qu'en ligne. C'est pourquoi le signalement dans le catalogue général des ressources électroniques pour lesquelles les accès sont pérennes ainsi que des documents de la filière ADDN est d'ores et déjà identifié comme une priorité.

Les instructions à conduire en 2016 pour répondre à ces enjeux porteront sur des solutions techniques comme le protocole OpenURL aussi bien que sur l'intégration de données depuis des sources externes.

Un signalement pertinent est l'une des clés du développement de l'accès à ces ressources.

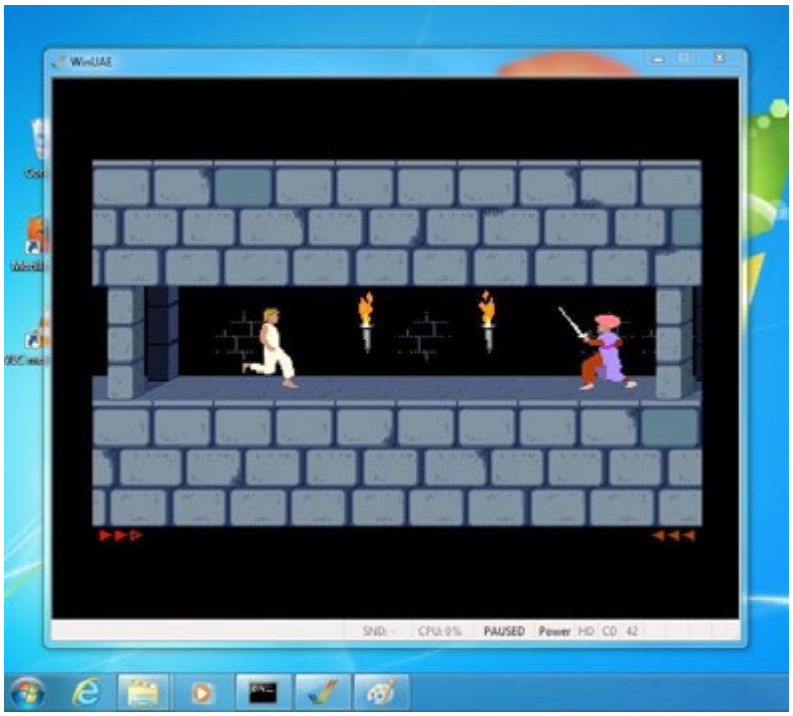
02.E Garantir une consultation pérenne des documents audiovisuels

L'ÉMULATION DES DOCUMENTS ÉLECTRONIQUES

À l'issue de la numérisation (voir axe 1 fiche B), certains documents demandent une intervention supplémentaire pour être pleinement lisibles et consultables. Les documents multimédias et interactifs, et plus encore les logiciels et les jeux vidéo, protégés par des procédés anti-copie, nécessitent généralement une intervention spécifique. Il a fallu développer un système propre pour réaliser l'émulation de ces documents anciens sur des postes informatiques récents, ce qui revient à simuler le fonctionnement des appareils de lecture et des systèmes d'exploitation d'époque.
Un programme pour copier et émuler les environnements logiciels d'origine, système par système, console par console, est ainsi en cours depuis le milieu des années 2000 : les documents pour plateformes Thomson (jeux, et surtout documents pédagogiques en lien avec le plan Informatique pour tous de 1985, pour TO7 et MO5) ont pu être traités, de même que les jeux sur disquettes des années 1980 pour les plateformes Amiga et Atari et la totalité des disquettes PC. Un nouveau chantier lancé en 2014 concerne les cédéroms.

LE SYSTÈME AUDIOVISUEL

Le système audiovisuel a été créé en 1996 pour l'archivage et la consultation des documents audiovisuels. Actuellement doté d'un budget annuel de 450 000 euros et maintenu au quotidien par un service de 6 personnes (ingénieurs, administrateur des systèmes et opérateur technique), le système est progressivement mis à niveau afin de répondre aux besoins de stockage de fichiers de plus en plus lourds et de plus en plus diversifiés. Cette infrastructure se double d'une interface évolutive de consultation des collections, régulièrement adaptée en fonction des besoins et des usages du public. Le système audiovisuel conserve actuellement plus de 813 000 documents correspondants à 1,7 pétaoctet. Actuellement, 80 % des consultations se font via cette infrastructure dédiée à partir de l'exemplaire numérique.



Émulation du jeu vidéo Prince of Persia (disquette, plateforme Amiga, 1990).

L'accès aux documents audiovisuels doit trouver son articulation avec l'Institut national de l'audiovisuel (INA), dont les collections sont consultables à la BnF, sur le site François-Mitterrand (offre grand public en Haut-de-jardin, salle de la presse et des médias – dépôt légal des radios et télévisions françaises en Rez-de-jardin à l'INAthèque).
L'intégration des documents audiovisuels numérisés de la BnF dans SPAR représente également un défi à relever dans les années à venir. Le système a été conçu pour les accueillir et en détient déjà une petite part, correspondant aux numérisations récentes effectuées dans le cadre de partenariats. Une migration massive devra tôt ou tard être envisagée : il s'agit là d'un projet complexe et de grande ampleur susceptible de durer plusieurs années.

L'intégration des documents audiovisuels numérisés de la BnF dans SPAR représente un défi à relever dans les années à venir.

02.F
La transition bibliographique

En France, les grands acteurs publics en charge de la normalisation documentaire se sont engagés conjointement, sous l'égide de l'AFNOR et à travers le programme national de la Transition bibliographique (transition-bibliographique.fr), dans un vaste chantier de révision des normes et consignes de catalogage courant, tout en assurant le traitement rétrospectif des données historiques afin de les mettre en conformité avec ce nouveau modèle.

UN NOUVEAU CODE DE CATALOGAGE

Ces travaux portent sur la rédaction d'un nouveau code de catalogage (RDA.FR) qui vise à adapter le code international RDA (*Resource Description and Access*) en vue de son adoption progressive. Ils s'accompagnent d'actions de formation et de sensibilisation des professionnels comme des éditeurs de systèmes d'information dans le but de faciliter l'accompagnement au changement sur une période évaluée à dix ans.

L'ENJEU DU WEB SÉMANTIQUE

Dans un contexte de croissance continue de la production, il est indispensable d'exploiter les possibilités offertes par les technologies du web sémantique. Celles-ci permettront de partager les données non plus en les répliquant ou en les dérivant, mais en les liant les unes aux autres. De manière à alléger ses coûts de catalogage, la BnF travaillera davantage en réseau afin de se concentrer sur sa valeur ajoutée propre. Il s'agit là d'une profonde mutation culturelle. L'intégration d'un nombre croissant de métadonnées issues directement des éditeurs et des distributeurs, ou encore la dérivation de notices internationales lors d'acquisitions de monographies étrangères, sont des exemples emblématiques de cette ouverture. C'est aussi la voie qu'ont retenue conjointement l'ABES et la BnF, qui souhaitent bâtir ensemble des référentiels nationaux d'autorité.

Tb
Transition bibliographique
FRBRiser les catalogues et adopter de nouvelles règles de catalogage dérivées de RDA

Dans un contexte de croissance continue de la production, il est indispensable d'exploiter les possibilités offertes par les technologies du web sémantique.

L'identification pérenne et standardisée des ressources constituera un autre point clé pour l'organisation et le partage de l'information. La BnF continuera à s'investir fortement dans la maintenance et dans l'évolution d'identifiants historiques (tels l'ISSN ou l'identifiant ARK – *Archival Resource Key*), tandis que de nouveaux chantiers s'engageront, au premier rang desquels celui de l'ISNI.

REFONDRE LES OUTILS DE PRODUCTION

En 2016 s'engage une première étude visant la refonte des outils de production, qui commencera par le prototypage d'un fichier national des données d'autorité en partenariat avec d'autres opérateurs publics, l'ABES au premier chef.
Bien que déterminées et irréversibles, ces avancées s'effectueront pas à pas. La BnF n'a pas souhaité précipiter une révolution du catalogage ; elle a préféré démontrer les bénéfices du changement tout en préparant par étapes progressives la transition bibliographique.

ISNI : un identifiant pour les personnes

La BnF a été la première bibliothèque au monde à intégrer aux notices d'autorité de son catalogue général des identifiants noms de personnes et de collectivités conformes à la norme ISNI (*International Standard Name Identifier*). Près d'un million d'ISNI ont ainsi été chargés à l'été 2013 et sont, depuis 2014, récupérables dans les produits bibliographiques et via data.bnf.fr.
La BnF est devenue agence d'enregistrement ISNI en 2014, se dotant ainsi d'un levier très concret de collaboration avec une pluralité d'acteurs : l'interprofession du livre (Dilicom et la Commission interprofessionnelle du livre), l'ABES, ou encore les Archives nationales. Très prometteuses, ces perspectives de collaboration pourraient préfigurer un dispositif national de gouvernance des référentiels auteurs et collectivités, sous la forme d'une fédération nationale d'opérateurs publics dont la BnF serait la tête de pont.

BnF

03

INNOVER AU SERVICE DES PUBLICS

3,5 millions
de documents dans Gallica

3,8 millions
de documents dans Gallica intra muros

15,8 millions
de visites de Gallica en 2015
correspondant à 330 millions de pages vues

+ de 170 000
abonnés au compte Twitter officiel de la BnF

+ de 110 000
abonnés à la page Facebook de Gallica

13 000
commandes de reproductions par an

Gallica figure parmi les premières bibliothèques numériques du monde. L'audience et la satisfaction des utilisateurs sont au rendez-vous ; Gallica et data.bnf.fr ont d'ailleurs obtenu en 2013 le prix de l'innovation décerné par les bibliothèques de l'université de Stanford. Il n'en reste pas moins que le défi du numérique se pose aussi à la BnF dans le domaine de sa mission de diffusion et de mise en valeur des connaissances. Il interroge la relation entre les services offerts sur place et ceux qui se déclinent en ligne, conduit à repenser les modalités de la médiation et de l'accompagnement à travers de nouveaux outils, invite à conquérir de nouveaux territoires, notamment au travers des réseaux sociaux.
En 2015, la BnF a lancé un chantier de dynamisation de sa politique des publics, avec pour objectif d'aller au-devant de ses différents publics, dans un esprit d'ouverture et de partage, afin de leur offrir une expérience positive, enrichissante et adaptée. Cette stratégie se décline dans le numérique, sur place comme à distance.

ZOOM SUR QUELQUES POINTS FORTS

03.A_ Le site www.bnf.fr
03.B_ L'éditorialisation et la médiation des contenus de Gallica
03.C_ Un écosystème de communication numérique
03.D_ Les outils en ligne d'aide à la recherche
03.E_ Diffusion et valorisation commerciale des collections numériques
03.F_ Les publications scientifiques

MOYENS, OUTILS ET MÉTHODES

03.G_ Les profils et les attentes des publics de Gallica
03.H_ Les canaux de la politique culturelle numérique
03.I_ L'accessibilité numérique

AXES DE DÉVELOPPEMENT

03.J_ Dissémination des ressources de Gallica et stratégie des API
03.K_ Développer les usages grand public de Gallica
03.L_ L'éducation artistique et culturelle
03.M_ Le prêt numérique
03.N_ L'innovation pour la consultation de la presse
03.O_ Le déploiement du Wi-Fi

Un Gallica, des Gallica

Jusqu'en 2012, Gallica se définissait comme une bibliothèque numérique unique donnant accès aux mêmes documents et proposant les mêmes fonctionnalités à tous les internautes, qu'ils se connectent à distance ou dans l'enceinte de la Bibliothèque. Début 2012, l'ouverture au public de Gallica intra muros a étendu l'horizon de la bibliothèque numérique en rendant accessibles dans les emprises de la BnF les contenus numérisés encore protégés au titre de la propriété intellectuelle. Gallica intra muros devient ainsi la plateforme principale de diffusion des contenus numériques, d'autant qu'elle a vocation à accueillir demain une partie des acquisitions et dons de documents numériques (ADDN – voir axe 5 fiche I) ainsi que les livres numériques entrés par dépôt légal. Parallèlement, la mise en place d'une plateforme de test Gallica Labs en 2012, le lancement des applications mobiles (iOS en 2012, Android en 2013), du site web mobile gallica.bnf.fr/m en 2013, et des premiers sites en marque blanche (Numistral en 2013, La Grande Collecte en 2014) ont rendu plus divers encore les accès aux fonds numérisés, qu'il s'agisse des contenus ou des fonctionnalités proposées dans ces différentes interfaces.

Une véritable communauté, les Gallicanautes, émerge entre érudition et créativité.

Aujourd'hui, Gallica se décline en une offre diversifiée, capable de s'adapter aux usages en mobilité sur tous types de terminaux, et répondant aussi bien aux attentes des lecteurs en salles que des internautes du monde entier.

Portée par une stratégie très active sur les réseaux sociaux, la bibliothèque numérique s'exporte aussi sur son blogue, dans sa lettre d'information, sur Facebook, Twitter et Pinterest, dans des interfaces de programmation (API) comme IIIF (*International Image Interoperability Framework*), avec pour ambition de toucher un public toujours plus large et divers, du chercheur le plus spécialisé au curieux venu flâner dans l'immense richesse des collections numérisées. Une véritable communauté, les Gallicanautes, émerge ainsi, entre érudition et créativité.

Vers la salle de lecture numérique, partout et pour tous

Dans les salles de lecture de la BnF, l'enjeu est de proposer au lecteur un espace de travail personnalisable, qui lui donne conjointement accès aux documents physiques et aux ressources numériques tout en lui offrant un environnement optimal pour la conduite de ses recherches. L'équipement en accès internet par câble et par Wi-Fi, permettra aux chercheurs d'accéder aisément aux services libres et ouverts (site www.bnf.fr, catalogues et Gallica) ainsi qu'aux ressources numériques protégées par le droit d'auteur (archives de l'internet, ouvrages protégés de Gallica intra muros, ressources numériques acquises), et ce, en utilisant leur propre matériel. Lancé en 2014, le portail numérique AVEC (Apportez votre équipement personnel de communication, BYOD en anglais) entend satisfaire cette exigence, en donnant accès, sur les terminaux personnels des lecteurs, à l'ensemble des documents numériques disponibles en bibliothèque de recherche, y compris ceux encore protégés au titre de la propriété intellectuelle.

Les chercheurs et les lecteurs sont appelés à devenir de plus en plus acteurs, préfigurant une économie du savoir plus collaborative.

Au-delà des logiciels de bureautique (Word, Excel, PowerPoint) déjà disponibles, les ordinateurs de la Bibliothèque proposent également des outils logiciels spécialisés, comme c'est le cas au sein du studio de création audiovisuelle en Haut-de-jardin qui propose aux lecteurs de s'initier au maniement du son et de la vidéo, et à la modélisation 3D. Cette panoplie complète un ensemble de dispositifs déployés de longue date – l'offre audiovisuelle du Haut-de-jardin (site François-Mitterrand) est entièrement numérique depuis 2004 – qui permettent d'accéder à des contenus spécifiques (jeux vidéo, documents multimédias…), parfois finement éditorialisés dans cette perspective. Reste à accompagner les usagers dans l'appropriation des outils et des ressources mis à leur disposition : des ateliers autour des fonds documentaires comme ceux du PRISME (Pôle de ressources et d'information sur le monde de l'entreprise) sont proposés aux personnes en recherche d'emploi ou souhaitant créer leur activité. Ces initiatives s'étendront à l'avenir à d'autres logiciels utilisés pour manipuler des ressources numériques de nature variée par les chercheurs aussi bien que le grand public (NASA World Wind, MARBLE, logiciels de reconnaissance de forme, de retouche d'images, de confrontation d'images, visualiseur d'éléments 3D, design 3D, etc.).

Les outils qui seront mis à disposition des chercheurs dans les années à venir devront leur permettre d'appréhender de manière plus aisée, grâce aux technologies des mégadonnées (ou *big data*), les masses de données importantes produites ou collectées par la BnF : reconnaissance d'entités nommées, moteurs sémantiques, etc. Les chercheurs et les lecteurs sont appelés à devenir de plus en plus acteurs de la consultation et de l'indexation des documents, préfigurant une économie du savoir plus collaborative. Des outils interactifs comme la plateforme expérimentale Correct, permettant de corriger l'OCR des documents de Gallica, sont autant de préfigurations concrètes de ce bouleversement des pratiques.

Une médiation ambitieuse et plurielle

Levier privilégié de la diversification des publics, la médiation a pris une importance croissante dans une institution longtemps tournée vers un public de chercheurs et de lecteurs. Le projet culturel de la BnF (expositions, offre pédagogique, manifestations, visites) a ainsi été pensé, depuis vingt ans, comme un prolongement des salles de lecture, offrant au public un accès différent au patrimoine.

L'ouverture à des publics nouveaux – collégiens, lycéens et même, plus récemment, enfants au Centre national de la littérature pour la jeunesse – est l'une des conséquences remarquables de cette médiation qui est désormais au cœur des pratiques de la BnF. L'établissement se positionne, à travers les ressources numériques qu'il diffuse, comme l'un des acteurs majeurs de l'éducation artistique et culturelle.

La numérisation est devenue en peu de temps le principal levier d'accessibilité aux documents patrimoniaux.

La déclinaison, dans le monde numérique, de cette stratégie d'ouverture prend des formes diverses. Sur Gallica, des corpus numériques correspondant à un fonds ou à une sélection d'ouvrages de référence sont constitués par les départements de collections. D'autres produits sont créés à l'attention de publics spécifiques dans le cadre de collaborations et de partenariats avec des institutions extérieures, telles les rubriques Classes (classes.bnf.fr) ou Expositions (expositions.bnf.fr). S'adressant autant aux élèves et apprentis qu'aux enseignants, cette offre à visée pédagogique a aussi pour ambition, par d'autres intermédiaires, de toucher le public qui n'ose pas ou n'imagine pas franchir la porte d'un lieu culturel.

Des portails ont été développés dans le cadre de programmes internationaux de coopération numérique, permettant de créer des ensembles thématiques ou de reconstituer des ensembles dispersés, comme ce fut le cas avec le projet Europeana Regia, programme de numérisation de documents royaux européens. De nouveaux usages de ces fonds, entre autres à des fins de recherche, peuvent ainsi émerger.

Avec une offre abondante et variée, d'une part, et, d'autre part, l'introduction d'outils améliorant la qualité de lecture au bénéfice, notamment, des personnes en situation de handicap visuel, la numérisation est devenue en peu de temps le principal levier d'accessibilité aux documents patrimoniaux.

De nouveaux territoires ont enfin été conquis, en particulier grâce aux objets nomades – tablettes, téléphones, etc. – qui permettent à chacun de devenir médiateur de contenus. L'ambition de développer les usages grand public de Gallica s'inscrit pleinement dans la politique de développement des publics de l'établissement à partir de 2016.

03.A Le site www.bnf.fr

Au cœur de l'écosystème de communication numérique, www.bnf.fr constitue le portail commun par lequel s'échange un volume considérable d'informations et de services, adressés à des publics eux-mêmes nombreux et variés.

LE PORTAIL

Il présente la bibliothèque et donne accès à l'ensemble de son offre : outils de bibliothéconomie, ressources professionnelles, programmation culturelle et événementielle ou bien encore informations pratiques, autant de contenus qui bénéficient d'une visibilité forte via un site identifié par les usagers, et dont la refonte constitue un enjeu majeur dans les mois à venir.
Ce portail met en relief les sous-domaines de www.bnf.fr qui donnent accès aux services fondamentaux de la bibliothèque (catalogue, Gallica…) ainsi qu'à des services spécifiques (réservation de places et de documents).

LE SITE MOBILE

Mise en ligne en octobre 2013, la version du site web pour mobiles présente une architecture et un design qui lui sont propres, tout en utilisant les mêmes contenus que le portail, lesquels ont été réorganisés en fonction de leur utilité en situation de mobilité. Depuis avril 2015, cette version inclut l'espace personnel qui permet la réservation de places et de documents.

RETOUR SUR L'ÉTUDE DE SATISFACTION

Une étude qualitative menée au printemps 2015 par l'agence Creative Works a mis en lumière les points de satisfaction et les critiques des utilisateurs à l'égard du site. Cette étude permet de mieux cerner la perception actuelle du site et de sa version mobile. Ses conclusions révèlent des pistes d'amélioration du site www.bnf.fr, mais aussi la relation fidèle et passionnée des interviewés avec l'institution.
Les avis sont très positifs sur la version mobile, qui offre souplesse, fiabilité technique et gain de temps.

LA REFONTE DU SITE : UN CHANTIER TRANSVERSE

Une refonte globale du site a été lancée à l'automne 2015, avec pour objectif la mise en ligne d'une première version à la fin de l'année 2016. Mobilisant de façon transverse un grand nombre d'acteurs, et permettant à l'institution de repenser sa relation à ses publics, le projet sous-tend des enjeux majeurs : d'ordre technique (changement de CMS – adoption de Drupal – et migration des contenus existants), méthodologiques (choix de la méthode dite « agile » Scrum - voir axe 6 fiche D -, déjà utilisée à la BnF) et éditoriaux (au nom de la lisibilité, de la clarté et de la mise en avant des services et des ressources).

Une refonte globale du site a été lancée à l'automne 2015.

L'espace personnel du site mobile et la page d'accueil de www.bnf.fr

03.B
L'éditorialisation et la médiation des contenus de Gallica

Dans l'enquête sur les usages de Gallica menée en 2011, un usager déclarait : « Il faudrait pouvoir aider les (nouveaux) utilisateurs à la manière dont les bibliothécaires guident une recherche dans une bibliothèque physique ». Afin de répondre à cette attente, un mode d'accès didactique aux collections de Gallica a été créé sous la forme de 306 pages web de médiation numérique regroupées en 25 corpus.

UNE MÉDIATION NUMÉRIQUE PLUS VISIBLE

Conçues au moyen d'un système de gestion de contenus (Drupal), les pages de médiation numérique, en ligne depuis 2014, sont particulièrement mises en avant dans la nouvelle version de Gallica (2015) : valorisées dès la page d'accueil, elles sont accessibles depuis n'importe quelle page du site via le bouton « Collections ». Par leur intermédiaire, les voies d'accès sont démultipliées : un accès thématique et un accès géographique complètent l'accès par type de document, qui est lui-même enrichi par les catégories « objets » et « vidéos ».
À partir de 2016, un enrichissement des gabarits existants sera entrepris en fonction des demandes des contributeurs – frise chronologique, carte interactive, etc. Une réflexion sera parallèlement engagée avec les partenaires de la BnF participant à Gallica afin de définir les modalités de leur contribution aux pages de médiation.

LES GALLICANAUTES, UNE COMMUNAUTÉ ACTIVE

À plus long terme, la mise en valeur des contenus de Gallica – à partir du site lui-même, le cas échéant – pourrait s'appuyer sur un périmètre élargi de contributeurs : au-delà des agents de la BnF et de ses partenaires institutionnels, les Gallicanautes eux-mêmes pourraient être associés à l'enrichissement et à la création de contenus éditorialisés, suivant des procédures et à l'aide d'outils à imaginer. Dans cette perspective, l'ajout de liens vers des documents de Gallica dans les pages de Wikipédia représente un enjeu de visibilité.

Les Essentiels de la littérature

Au titre de son offre de médiation numérique, la BnF développe les Essentiels de la littérature, un site spécifique visant à offrir à un large public (niveau lycée) un accès simple aux œuvres essentielles de la littérature. Celles-ci sont disponibles via Gallica, à travers une interface éditorialisée proposant des modules sur les auteurs, les œuvres, et des parcours thématiques.

Un prototype consacré à Montesquieu a été mis en ligne en octobre 2014, avec deux modules oeuvre, consacrés aux *Lettres persanes* et à *L'esprit des lois*, ainsi qu'un module consacré à l'esclavage, donnant un bon aperçu de ce projet de médiation numérique. Le projet s'est poursuivi en 2015, avec la mise en ligne du XVIIIe siècle, qui sera suivie par celle du XIXe siècle et du début du XXe siècle.

03.C Un écosystème de communication numérique

Grâce à un site web institutionnel « portail » proposant de nombreux services et sous-sites à destination des professionnels ou du grand public, la BnF bénéficie d'une place d'acteur majeur du numérique en France et dans le monde. Elle développe et décline cette identité forte via des canaux de médias sociaux diversifiés. Blogues, carnets de recherche sur plateforme collaborative, lettres d'information, chaînes vidéos et comptes de réseaux sociaux : autant de passerelles tissées entre la bibliothèque et ses publics.

LE RÔLE PRÉPONDÉRANT DES RÉSEAUX SOCIAUX

Dans cet écosystème, les réseaux sociaux numériques occupent une place privilégiée. Facebook, Twitter, mais aussi Pinterest, Instagram, YouTube, Dailymotion, iTunesUniversity permettent d'ancrer la présence institutionnelle de la BnF via des comptes dits « officiels ». Ces canaux de communication généraliste, axés sur la programmation et la vie de l'établissement mais aussi ses collections et ses enjeux patrimoniaux, contribuent à une image dynamique de la bibliothèque, en prise directe avec l'actualité. D'autres comptes de réseaux sociaux mettent en avant la richesse des fonds et la diversité des offres de services via des comptes dédiés (comptes liés à la bibliothèque numérique Gallica, blogue Ad vivum pour le département des estampes et de la photographie, par exemple) qui favorisent les discussions relatives aux activités d'un département et l'approfondissement des liens avec les communautés de chercheurs et spécialistes.

Ces canaux de communication généraliste contribuent à une image dynamique de la bibliothèque, en prise directe avec l'actualité.

LES TROIS AXES D'UNE STRATÉGIE À LONG TERME

Pour asseoir et affirmer cette nouvelle BnF, virtuelle, vivante, trois pistes se dessinent.

Premièrement, renforcer sa présence en ligne par une communication numérique ambitieuse, plus coordonnée et au diapason des nouvelles pratiques. Plusieurs mesures y concourent : deux formations « Communiquer sur les réseaux sociaux » et « Écrire pour le web » proposées aux agents, une charte d'utilisation des réseaux sociaux et la mise en place d'outils, tant pour la définition de lignes éditoriales que pour la planification des publications.

La deuxième piste repose sur la fidélisation des communautés d'internautes et l'approfondissement du sentiment d'appartenance des usagers de la Bibliothèque (invitations permettant aux publics – Gallicanautes, etc. – de devenir des ambassadeurs de la BnF, recherche de relais et de communautés d'internautes plus éloignées ou s'intéressant à des sujets rares).

Le troisième axe consiste à élargir les actions de veille et de gestion de l'e-réputation. Prise de recul, analyse de la performance, meilleure compréhension des attentes des publics, gestion de crise : tels sont les enjeux d'une veille élargie et organisée sur les présences en ligne de la Bibliothèque en dehors de ses propres canaux. Dans ce domaine, l'établissement peut s'appuyer sur un atout considérable : la crédibilité que le public accorde à son action de valorisation patrimoniale.

Les Gallicanautes

Les internautes qui utilisent Gallica et participent activement à la diffusion de ses documents sur le web (les Gallicanautes) ont une place au sein de la bibliothèque numérique qui leur consacre une page spécifique dans sa nouvelle version « Du côté des Gallicanautes ». Du réemploi spécifique au détournement récréatif des collections les plus diverses (manuscrits médiévaux aussi bien que journaux du XIXe siècle), les réutilisations des Gallicanautes forment une palette riche et variée : coloriages vintage, découpages et collages de globes, détournement d'enluminures, création de GIF animés, réalisation de (vieilles) recettes de cuisine, illustrations de monuments visités au cours d'un voyage… Ainsi, Thomas Henry, gallicanaute passionné par l'histoire du disque et auteur du blogue *Ceints de bakélite,* a monté le projet « Disquaires de Paris » et créé un site web dédié à l'histoire des disquaires parisiens où il propose une carte interactive de Paris donnant accès à des ressources numérisées souvent issues de Gallica.
De son côté, Johanna Daniel, dite « Peccadille », a créé en 2012 le blogue *Orion en aéroplane* pour « partager les belles choses que ses études en histoire de l'art lui faisaient découvrir ». Elle y a notamment proposé d'utiliser l'application Gallica pour tablettes et téléphones portables afin de retracer l'évolution des monuments de Paris.

03.D
Les outils en ligne d'aide à la recherche

SINDBAD : APPORTER DES RÉPONSES FACILEMENT CONSULTABLES

Afin de traiter avec efficacité les nombreuses questions qui lui parviennent, la BnF a créé un service dédié, SINDBAD (Service d'information des bibliothécaires à distance). Adressées par messagerie instantanée, par téléphone ou par formulaire web, les interrogations des usagers y reçoivent une réponse rapide, en direct ou par courriel.

Plus de 10 000 questions y sont traitées chaque année. 1 800 des réponses apportées à ce jour ont été archivées dans « La base de connaissances », hébergée sur le site « Question point ». Il est prévu d'intégrer ces questions-réponses dans les pages du site www.bnf.fr, ce qui permettrait leur moissonnage par les moteurs de recherche. L'internaute consultant la base pourrait, par ailleurs, trouver réponse à ses questions sans quitter le site de la BnF.

LE SITE « CHERCHER & TROUVER »

Ce dispositif d'aide à la recherche est complété depuis janvier 2014 par le site « Chercher & trouver », où sont proposés une vingtaine de courts mémos de recherche indiquant comment trouver un article de presse, l'ayant droit d'un film, des récits de voyage… Chaque mémo précise quelles sources utiliser et comment les exploiter. De nouveaux mémos viendront enrichir le site afin d'élargir les domaines couverts.

DES PORTAILS D'AIDE À LA RECHERCHE

Fin 2012 ont été mis en ligne quatre portails produits avec le logiciel LibGuides, qui prolongent les quatre centres de ressources thématiques créés dans les salles du Haut-de-jardin : Développement durable, Europe, Francophonie, Sociétés en débats. En 2014, une communication régulière autour de ces portails sur www.bnf.fr, la page Facebook de la BnF et le compte Twitter @ActuBnF a entraîné une augmentation de 36 % de leur fréquentation.

Sur le plan des contenus, les portails sont appelés à accueillir plus de mini-dossiers, articulés autour de sujets précis. À titre d'exemple, la partie « Débats » du portail Europe, l'une des plus consultées, a été étoffée en 2014-2015 de nouveaux dossiers (« La Turquie dans l'Union européenne ? », « L'UE souffre-t-elle d'un déficit démocratique ? », « TTIP, piège ou opportunité pour l'Europe ? »…), assortis de conseils de lecture, sites web et vidéos permettant d'appréhender les différents points de vue sur ces questions.

Depuis 2015, la BnF offre également un site de ressources et d'information sur les métiers du livre, le portail Métiers du livre.

Au sein de ces portails, plusieurs conseils de lecture, sites web et vidéos permettent d'appréhender les différents points de vue sur des questions d'actualité.

Exemple de portail thématique : le portail Europe

03.E Diffusion et valorisation commerciale des collections numériques

Avec 13 000 commandes par an, le service de numérisation à la demande assure la reproduction payante d'environ 25 000 documents numériques – 1,2 million de pages reproduites – auprès de 6 500 clients, français et étrangers.

Ce service enrichit l'offre de communication aux lecteurs en leur permettant d'accéder, via une prestation commerciale, à des documents non encore numérisés. Il valorise également les collections de la Bibliothèque par la vente d'images numériques aux professionnels (éditeurs, musées, etc.). 1 500 projets commerciaux donnant lieu au paiement d'une redevance d'utilisation sont ainsi réalisés chaque année sur la base des fonds de la BnF : éditions, productions audiovisuelles, expositions, sites web, applications mobiles… L'ensemble de ces activités génère en moyenne 1,4 million d'euros de recettes annuelles pour l'établissement, avec un objectif d'augmentation de 200 000 euros par an d'ici 2018. Cette vente s'effectue par le biais de la Banque d'images, qui propose uniquement des documents iconographiques libres de droit numérisés à la demande de ses clients (environ 200 000 images).

ADAPTER LA GAMME DE SERVICES

Dans un marché des contenus numériques extrêmement concurrentiel, la mise en œuvre d'une stratégie ambitieuse de diffusion et de valorisation commerciales exige que la gamme de services s'adapte régulièrement aux attentes des clients. Cet impératif doit conduire à faire évoluer constamment l'offre de reproduction, avec une hausse des standards et de la qualité de numérisation, une diversification des produits proposés, et la mise en place de tarifs adaptés.

Le développement de nouveaux outils de vente et de valorisation des contenus s'impose également - vente en ligne de reproductions, nouvelle banque d'images BnF, simplification des parcours clients, accès et médiation vers les contenus, etc. L'impact de ces améliorations sur les clients et leurs pratiques sera évalué à l'aide d'un baromètre de satisfaction administré en ligne et des résultats commerciaux.

PERSPECTIVES LIÉES AUX TECHNIQUES DE RECONNAISSANCE VISUELLE

Pour améliorer le signalement et la visibilité en ligne de son patrimoine iconographique numérique, la BnF mène, par ailleurs, avec le laboratoire ETIS du CNRS, un projet scientifique intitulé ASAP (Annotation semi-automatique des images patrimoniales) dont l'objectif est de proposer de nouvelles solutions d'aide à la description des images numérisées. Reposant sur les algorithmes les plus récents de reconnaissance visuelle par similarité du contenu, le travail portera sur la conception de fonctions d'indexation automatique calculées à l'aide d'images de la BnF déjà annotées.

La possibilité de diffuser images et métadonnées en masse devrait faciliter la constitution d'un réseau de distribution via des agences photos étrangères.

Banque d'images, version 2

Afin de soutenir le développement des recettes, une nouvelle banque d'images, en cours de conception, présentera en un seul et même endroit une grande partie des documents iconographiques de la Bibliothèque et intégrera de nombreux fonds sous droits aujourd'hui sous-exploités.

Les fonctionnalités d'achat et de téléchargement intégrées permettront d'automatiser la vente et de gérer les redevances d'utilisation. La possibilité de diffuser images et métadonnées en masse devrait faciliter la constitution d'un réseau de distribution via des agences photos étrangères.

03.F
Les publications scientifiques

La BnF participe de manière volontariste à la politique nationale de libre accès (*open access*), qui encourage le partage des résultats de la recherche, leur diffusion et leur conservation. La mise en ligne de la bibliographie des publications des chercheurs de la BnF répond au premier chef à cette exigence : elle est effectuée à travers le site Production scientifique de la Bibliothèque nationale de France, tandis que le texte intégral des articles est diffusé sur le portail d'archives ouvertes de la BnF dans HAL (Hyper articles en ligne). Parallèlement, l'établissement s'attache à mettre en ligne et actualiser sa base des Actions nationales, internationales et de recherche (ANIR), recensant les programmes de recherche de la BnF.

L'EXCELLENCE SCIENTIFIQUE

À des fins de mise en valeur des collections, la BnF édite, par ailleurs, des catalogues raisonnés et des ouvrages scientifiques destinés aux chercheurs et aux professionnels des bibliothèques. Ces publications de référence sont le reflet de l'excellence scientifique de l'institution. Autour d'elles est construite une active politique d'édition électronique : après avoir lancé son offre sur le portail « OpenEdition Books », la BnF entend accroître la diffusion de ses productions scientifiques via cette plateforme qui comptait 27 ouvrages en mars 2016.

UNE DIVERSITÉ DE PUBLICATIONS

Conçue pour mettre en lumière les collections et faire connaître le travail des conservateurs et des chercheurs de l'établissement, la *Revue de la Bibliothèque nationale de France* est intégralement disponible sur le portail Cairn.info avec une barrière mobile de 36 mois. *La Revue des livres pour enfants* explore quant à elle les thématiques à l'œuvre dans la littérature pour la jeunesse et témoigne de l'activité critique comme de la veille des professionnels du CNLJ (Centre national de la littérature pour la jeunesse). Elle est disponible sur son site internet avec une barrière mobile de 24 mois. La Bibliothèque veille par ailleurs à signaler ses manifestations scientifiques au sein de Calenda, le calendrier scientifique francophone des lettres et sciences humaines et sociales, et à publier des carnets de recherche sur la plateforme Hypothèses.org du Centre pour l'édition électronique ouverte (Cléo). Les outils de la communication scientifique directe sont pleinement exploités par la BnF, laquelle est allée jusqu'à constituer un groupe LinkedIn (« BnF – cercle des chercheurs »), dédié aux chercheurs associés et visant à faire vivre en ligne la communauté de recherche de la Bibliothèque.

DES SITES DÉDIÉS

Certains programmes conduits au sein de l'établissement aboutissent à la production de sites web. C'est notamment le cas de Mandragore, base iconographique du département des manuscrits et de la bibliothèque de l'Arsenal, proposant plus de 170 000 notices d'analyse d'œuvres, souvent accompagnées d'une image numérisée. C'est également le cas de Reliures de la Bibliothèque nationale de France, qui présente des reliures françaises à décor du XVIe siècle au début du XIXe siècle conservées à la Réserve des livres rares. La Bibliographie des éditions parisiennes du XVIe siècle est un autre de ces sites témoignant de la volonté de la Bibliothèque d'affirmer son rôle actif dans la production des savoirs.

Une active politique de publications en ligne permet de faire connaître le travail des conservateurs et des chercheurs de l'établissement.

03.G Les profils et les attentes des publics de Gallica

Face à la multiplicité des pratiques numériques, des supports, des environnements web, élaborer une vision globale des usages en ligne et suivre leur évolution constituent un véritable défi. La collecte d'informations sur les utilisateurs de l'interface Gallica participe également de cette vaste entreprise de compréhension des usages.

ANALYSER LES USAGES

Une première enquête de grande envergure, conduite en 2011 auprès de 3 800 Gallicanautes, a permis de vérifier le poids des usages documentaires dits « traditionnels » : consultation de Gallica dans le cadre de recherches précises, professionnelles ou amateurs. Les usages de simple curiosité ou de divertissement demeuraient quant à eux l'exception. Ces résultats ont été complétés par des enquêtes qualitatives grâce auxquelles les parcours de chercheurs professionnels et amateurs ont pu être reconstitués autour d'un thème ou d'un corps : manuscrits en ligne, fonds spécialisés en sciences et techniques, patrimoine de la Grande Guerre.

Le profil des Gallicanautes

Selon l'étude de 2011, les Gallicanautes se connectent majoritairement depuis la France, même s'ils sont nombreux à utiliser Gallica depuis l'Italie, les États-Unis, le Canada, la Belgique, l'Espagne ou encore le Brésil. Lecteurs plus intensifs et plus diplômés que la moyenne des Français, ils sont aussi plus âgés que ceux des salles de lecture de la BnF, et plus fréquemment équipés d'une liseuse et d'une tablette.

Si 95 % des utilisateurs se déclarent satisfaits de Gallica, leur degré d'exigence demeure élevé, avec des attentes aiguisées par un environnement fortement concurrentiel (Google Books, etc.) : accroissement des fonds numérisés, facilitation de la navigation, amélioration du moteur de recherche et des outils d'exploitation des documents.

MESURER LES ÉVOLUTIONS

Une grande étude est lancée en 2016 afin de mesurer les évolutions, depuis l'étude de 2011, de la composition du public de Gallica et de ses usages et afin d'évaluer si les mutations importantes que la bibliothèque numérique a connues ces dernières années ont permis une diversification des publics et une évolution de leurs pratiques. Elle vise, en outre, à appréhender les logiques actuelles de parcours (entre différentes ressources documentaires, différents supports, différentes interfaces).

LE BIBLI-LAB

La BnF a aujourd'hui toute légitimité pour se positionner comme observateur privilégié et expert des usages du patrimoine numérisé : un positionnement tant institutionnel que scientifique, conforté par la convention de partenariat signée en janvier 2013 avec l'école Télécom ParisTech. D'une durée de cinq ans, ce partenariat s'est d'ores et déjà concrétisé par la création du Bibli-Lab, Laboratoire d'étude du patrimoine en ligne des bibliothèques. Il se décline en conventions d'application et en projets financés dans le cadre d'appels à projets nationaux et internationaux. Un premier programme de recherche a notamment permis d'identifier les enjeux théoriques et méthodologiques soulevés par l'analyse des usages de Gallica dans toutes ses dimensions de bibliothèque numérique (interface, réservoir de documents et services associés). Des préconisations ont pu être formulées, invitant notamment à évaluer les usages par ensembles documentaires, en lien avec les enjeux stratégiques de la BnF en matière de numérisation et de mise en valeur des collections numériques.

Un deuxième programme porte sur la circulation à travers le web des documents numérisés autour de la Grande Guerre, avec la collaboration de la BDIC, dans le cadre du Labex « Les passés dans le présent » (voir axe 6 fiche G).

Le Bibli-Lab, laboratoire d'étude du patrimoine en ligne des bibliothèques, se positionne comme observateur privilégié des usages.

03.H Les canaux de la politique culturelle numérique

Déjà bien présente dans les pratiques de l'établissement, la politique culturelle numérique, qui combine l'usage des outils numériques sur place (les audioguides multimédias, par exemple) ou à distance, ne demande qu'à se développer.

SUR LE PLAN DE LA DIFFUSION

Il s'agit d'accélérer la circulation des contenus et des savoirs dans différentes communautés d'internautes, des plus généralistes aux plus savantes. Tous les enregistrements de conférences, de stages de formation, d'interventions publiques sont à disposition sur le site de la BnF et bénéficient également de sites et portails relais (Radio France, iTunes U, les sites des institutions ou universités avec lesquelles la BnF collabore). Les expositions en ligne et autres albums numériques sont également disponibles sur des portails associés comme le portail de l'histoire des arts, du ministère de la Culture, ou le portail Éduthèque du ministère de l'Éducation nationale. Les blogues chercheurs, qui constituent une partie de la médiation des collections – du type hypotheses.org – sont régulièrement repris et moissonnés par des moteurs de la communauté scientifique susceptibles d'accroître leur circulation.

SUR LE PLAN DE LA PRODUCTION

La BnF a lancé, en collaboration avec Orange, une collection de livres enrichis sur de grands titres de la littérature française. *Candide, Dom Juan, Au bonheur des dames*, autant de productions éditoriales autonomes qui constituent des terrains de réflexion nécessaires à l'élaboration d'une politique éditoriale numérique incluant notamment un volet commercial (en téléchargement ou en produits).
Attentive à l'évolution des modalités d'accès à la connaissance et désireuse de s'adapter aux pratiques des publics, en particulier de répondre à leurs attentes croissantes en matière d'autoformation, la BnF souhaite également développer un projet de cours en ligne ouvert à tous (MOOC).

Il s'agit d'accélérer la circulation des contenus et des savoirs dans différentes communautés d'internautes, des plus généralistes aux plus savantes.

Les différentes formes de médiation

La médiation prend des formes variées, s'adressant à des publics dont la Bibliothèque doit encore améliorer sa connaissance afin de mieux adapter son offre :

- **Médiation de formation :** elle aide les lecteurs à se repérer dans les catalogues ou les collections et s'adresse surtout aux étudiants insuffisamment formés à la recherche ;
- **Médiation d'initiation :** liée à la visite des bâtiments ou d'expositions, elle vise à faire comprendre ce qu'est une bibliothèque nationale ou un propos d'exposition. L'engouement grandissant qu'elle suscite invite à des modélisations numériques (bornes interactives…) ;
- **Médiation d'éducation :** dédiée au public enseignant et scolaire, elle rencontre un succès croissant. La part du numérique, qui y est importante, va en se développant ;
- **Médiation d'animation :** privilégiant le divertissement comme moyen d'accès, elle est adaptée aux publics des ateliers « enfants » et familiaux. Amenant à découvrir la BnF, elle a vocation à être renforcée ;
- **Médiation de sensibilisation :** elle est destinée à des publics « éloignés » par la langue ou par les pratiques culturelles. Dans ce cadre s'inscrit le projet Passerelles, site dédié aux apprentis du BTP. Il propose une exploration interactive de divers métiers, notamment à travers quelque 70 constructions représentatives de l'histoire et des techniques de leur époque.

03.I L'accessibilité numérique

De même que pour l'accueil des publics dans ses emprises physiques, la BnF place l'accessibilité à son offre numérique au premier plan de ses préoccupations, qu'il s'agisse de ses sites et services en ligne ou des contenus qu'elle propose : cela passe à la fois par une attention accrue au développement d'interfaces adaptées à tous les types de handicap (visuel, auditif, moteur, cognitif…) et par la création de collections dédiées.

FORMER LES ÉQUIPES INFORMATIQUES

La BnF a entrepris de développer en priorité les compétences des équipes en charge de la réalisation de la nouvelle version de Gallica et de la refonte du catalogue général, avant d'étendre cette sensibilisation à l'ensemble des équipes chargées du développement des services en ligne. Elle s'attache à suivre les recommandations du Référentiel général d'accessibilité pour les administrations (RGAA version 2.2.1), auxquelles le site web mobile de Gallica est déjà conforme.

DÉVELOPPER L'ACCESSIBILITÉ DES COLLECTIONS

Le second défi majeur pour la BnF consiste à proposer davantage de documents accessibles à un public en situation de handicap. L'exploitation des potentialités du format EPUB – qui se développera dans les collections grâce au dépôt légal du livre numérique – va dans ce sens (voir axe 6 fiche A et axe 1 fiche H) : ainsi, les 3 200 ouvrages de Gallica traités dans ce format ont été systématiquement produits au format accessible DAISY.
En raison de son rôle dans l'application de l'exception liée au handicap dans le code de la propriété intellectuelle via le dispositif PLATON (voir axe 5 fiche D), la BnF est en capacité de devenir, en lien avec les associations agréées qui prennent en charge l'adaptation et la diffusion des textes, l'un des acteurs majeurs de l'accessibilité au livre numérique, à la fois en contribuant à un meilleur signalement des œuvres adaptées et en œuvrant avec les éditeurs à l'inclusion du livre numérique dans le dispositif PLATON.

SUR PLACE : PERSONNALISER L'ACCUEIL

Le numérique décuple le potentiel de mise à disposition d'outils et de dispositifs accessibles pour mieux accueillir ces publics spécifiques dans les salles de lecture et les expositions : vidéoagrandisseurs, machines à lire, applications accessibles aux personnes en situation de handicap visuel ou auditif… Un système de communication par visio-interprétation en langue des signes et transcription instantanée de la parole, qui permet de communiquer sur place et à distance avec les publics sourds et malentendants, est ainsi opérationnel depuis fin 2015.

La BnF est en capacité de devenir l'un des acteurs majeurs de l'accessibilité au livre numérique.

Vidéoagrandisseur.

03.J Dissémination des ressources de Gallica et stratégie des API

La capacité de Gallica à disséminer ses ressources en facilitant leur référencement et leur circulation en ligne est cruciale pour accroître la visibilité du site sur le web. Trois grandes actions y concourent : le développement d'outils de partage permettant à l'usager de s'approprier les contenus et de les disséminer facilement ; l'introduction de dispositifs améliorant le référencement des ressources de Gallica dans les moteurs de recherche ; la mise en œuvre d'initiatives visant à l'établissement d'une présence de Gallica sur des sites tiers (partenaires et extérieurs).

DÉVELOPPER LES OUTILS DE PARTAGE ET RENFORCER LE RÉFÉRENCEMENT

En cohérence avec ces trois mécanismes, de nombreux outils ont été mis en place : liens pérennes, lecteur exportable, partage sur les réseaux sociaux… Ils font l'objet d'une amélioration constante. Le perfectionnement du référencement naturel des ressources de Gallica a également été engagé, par optimisation des pages web et multiplication des liens entrants (depuis les pages data.bnf.fr par exemple).
Gallica développe, par ailleurs, une politique de dissémination via des interfaces de programmation (API), à usage interne et pour ses partenaires dans un premier temps. Ce sont ces interfaces qui permettent l'exploitation des données de Gallica dans le site mobile et les applications pour téléphones portables et tablettes.

ACCROÎTRE LA VISIBILITÉ

Les métadonnées des ressources de Gallica sont diffusées depuis l'origine de manière standard et réutilisables via l'entrepôt OAI-Num. Elles sont incluses dans la stratégie de données ouvertes de la BnF (voir axe 2 fiche A). L'effort pour offrir à des partenaires un accès distant normalisé aux ressources numérisées s'est toutefois amplifié avec la mise en service de l'API *International Image Interoperability Framework* (IIIF), un protocole conçu par l'université de Stanford pour permettre la consultation d'images et de portions d'images sur des sites distants. Ce protocole est soutenu par un consortium dont la BnF est membre fondateur.
Ce type d'outils permet à des développeurs externes de s'approprier les ressources de Gallica au travers d'interfaces techniques faciles à prendre en main. Directement diffusées sur des sites partenaires, ces ressources voient dès lors leur audience se démultiplier.

Les métadonnées des ressources de Gallica sont incluses dans la stratégie de données ouvertes de la BnF.

Étude d'un manuscrit de Gallica via l'API IIIF sur le prototype Biblissima

03.K
Développer les usages grand public de Gallica

La BnF place parmi ses axes de travail prioritaires l'ambition d'accroître la dimension grand public de la bibliothèque numérique. Il s'agit désormais d'aller chercher des usages qui dépassent les cercles des étudiants, chercheurs, amateurs et passionnés qui constituent le public traditionnel de Gallica. Le développement des usages de loisir et de découverte est le levier qui permettra de faire apparaître Gallica dans le paysage quotidien des internautes.

CRÉER L'ÉVÉNEMENT

Diversifier l'offre de médiation et étendre les canaux de diffusion permettront de toucher un public plus large. L'approche éditorialisée des contenus dans la rubrique « Collections » de Gallica, complétée par des publications sur d'autres plateformes, comme les Essentiels de la littérature, participe de cette dynamique. Sur les réseaux sociaux, l'action de l'équipe des animateurs de communauté *(community managers)* de Gallica favorise la diffusion sur un mode collaboratif, en mettant en avant les découvertes et en organisant des événements interactifs et ludiques. En outre, se poursuit le travail de mise en valeur de cette ressource auprès des publics très divers accueillis dans le cadre d'ateliers pédagogiques.

IMPLIQUER LES PUBLICS

Parmi les projets à l'étude, le projet « Gallica Studio », proposera une nouvelle approche impliquant directement les publics dans la création à partir des contenus de Gallica, ainsi réappropriés et revalorisés. Dans une interface distincte de celle de la bibliothèque numérique, les internautes pourront importer des documents de Gallica afin de les corriger, de les enrichir, les géolocaliser, les annoter, ou encore réaliser des projets créatifs. Les documents enrichis pourraient ensuite, le cas échéant, être reversés dans Gallica.

ÉLABORER UNE NOUVELLE OFFRE

En enrichissant Gallica de publications récentes susceptibles d'intéresser de nouveaux lecteurs, le projet de prêt numérique participe d'une nouvelle offre de services pour un public élargi.

L'*ABC de Babar* par Jean de Brunhoff sur Gallica

Un enchantement pour petits et grands

De toutes époques et de toutes formes, les livres pour les enfants et la jeunesse sont extrêmement nombreux dans Gallica : imagiers, abécédaires, albums, contes, classiques illustrés, en passant par l'imagerie populaire, telle celle d'Épinal, ludique, comme le coloriage, ou pédagogique, avec les cartes. Nombre de ces documents se prêtent à des usages familiaux comme la lecture de contes ou le livre du soir, ou à la (re) découverte de textes universels de la mythologie ou de classiques comme Perrault, la comtesse de Ségur, Jules Verne… De tels usages demandent une ergonomie adaptée (interfaces tactiles, tablettes dédiées de type Vtech ou Gulli…) et des corpus cohérents. Proposer un univers Gallica pour les enfants à travers une interface web spécifique et une application mobile permettra à tous de découvrir une offre adaptée aux différentes tranches d'âges (0-3 ans, 3-7 ans, 7-12 ans), élaborée par les spécialistes de la culture enfantine du CNLJ (Centre national de la littérature pour la jeunesse). Les contenus pourraient adopter une forme de feuilleton, rappelant l'abonnement aux magazines pour enfants.

03.L L'éducation artistique et culturelle

UN SERVICE À VOCATION NATIONALE

À partir de 2015, la BnF étend sur le territoire national son champ d'intervention dans le domaine de l'éducation artistique et culturelle (EAC) : dans le cadre des pôles associés régionaux, une nouvelle offre de formation, coordonnée et structurée, est mise en place autour du patrimoine du livre, domaine d'excellence de l'EAC à la BnF. Elle couvre les ressources numériques, l'histoire du livre et de la presse, l'éducation aux médias et à l'information, l'image et ses techniques, la littérature de jeunesse. Il s'agit d'enrichir l'offre culturelle nationale et de contribuer à l'éducation culturelle et artistique, tout en renforçant la coopération de la BnF avec les territoires : deux objectifs majeurs du contrat de performance 2014-2016.

EXPOSITIONS ET JOURNÉES DE FORMATION

Cette action prend la forme d'une exposition diffusable dans les établissements sur le patrimoine du livre et la BnF.

Une journée de formation en région est proposée, en outre, aux différents acteurs concernés : enseignants, professeurs-documentalistes, bibliothécaires, chargés de médiation (musées, archives), animateurs du périscolaire, élus, directeurs des affaires culturelles, libraires ; inspecteurs de l'inspection générale des bibliothèques ; correspondants des Délégations académiques aux arts et à la culture (DAAC) des rectorats, inspecteurs de la vie scolaire ; inspecteurs d'académie, inspecteurs pédagogiques régionaux (collège et lycée), etc.

L'objectif de cette formation consiste à montrer la complémentarité d'une approche EAC reposant à la fois sur des ressources numériques et sur des supports physiques. Elle consiste en particulier à :

- **Diversifier les approches du patrimoine :** les documents patrimoniaux d'une Bibliothèque municipale à vocation régionale (BMVR) sont présentés, en mettant l'accent sur le lien entre document original et document numérique, et sur les métiers et techniques (rencontre avec les professionnels, rapports avec la culture scientifique, technique et industrielle).
- **Faire connaître l'offre de la BnF :** cette formation est également l'occasion de faire connaître à de nouveaux publics les ressources numériques de la BnF (Gallica ; data.bnf.fr ; CCFr…) et le portail classes.bnf.fr. Une mallette pédagogique (matériauthèque, clé USB, tutoriel…) leur donne une idée concrète des ateliers pédagogiques élaborés par la Bibliothèque.
- **Fédérer les acteurs d'un même territoire :** Enfin, la formation inclut la présentation de l'enquête menée par le service du livre et de la lecture du Ministère sur la mise en place de l'éducation artistique et culturelle en bibliothèques, ainsi que les modalités de mise en œuvre des parcours d'EAC à travers des exemples sur le territoire.

Enrichir l'offre culturelle nationale et contribuer à l'éducation culturelle et artistique sont deux des objectifs majeurs du contrat de performance 2014-2016.

03.M
Le prêt numérique

La libre circulation de l'information sur internet a bouleversé les usages des lecteurs : les pratiques documentaires ne sont plus limitées par les contraintes de l'espace (les salles de lecture des bibliothèques) ou du temps (les horaires d'ouverture). Les usagers des bibliothèques demandent de plus en plus une offre de lecture disponible à tout moment, en ligne et en mobilité, sur leur propre matériel, ordinateur portable, tablette ou téléphone portable, quel que soit le lieu où ils se trouvent. La transition numérique de la filière de l'édition s'accompagne ainsi d'une adaptation de l'offre de prêt en bibliothèque.

ENTRE MISSION DE LA BNF ET ENJEU DE SOCIÉTÉ

Si la BnF, dont la mission première est la conservation d'une collection à vocation patrimoniale, n'a jamais été une bibliothèque de prêt, Gallica joue en réalité un rôle comparable pour les titres du domaine public ou dont les droits ont été négociés, la consultation des ouvrages protégés par le droit d'auteur restant en revanche limitée aux emprises physiques de la Bibliothèque.
Quant aux ressources électroniques acquises, un accès distant est offert depuis 2009 aux seuls lecteurs disposant d'une carte annuelle recherche.

UNE NOUVELLE OPPORTUNITÉ

En 2014-2015, la BnF a entrepris de définir une nouvelle politique de développement des publics visant à dynamiser sa fréquentation aussi bien sur place qu'à distance. L'ambition de toucher un public plus large trouve naturellement son écho dans un projet de prêt numérique, qui permettrait de s'adresser aux usagers sur tout le territoire. Le service de prêt numérique imaginé par la BnF pourrait s'adresser demain à tout lecteur inscrit annuellement à la BnF (Haut-de-jardin, Rez-de-jardin, Richelieu et autres sites).

UNE POLITIQUE SPÉCIFIQUE

La BnF ne peut concevoir la politique documentaire de son offre de prêt numérique qu'en recherchant la complémentarité avec celle des bibliothèques publiques et universitaires, qui proposent déjà des services de prêt ou d'accès distant à des livres numériques. Cette politique documentaire serait construite dans la continuité de celle des acquisitions de la bibliothèque du Haut-de-jardin et en conformité avec sa vocation de bibliothèque patrimoniale : des ouvrages de référence, encyclopédiques et généralistes, et dont la durée d'usage dépasse l'actualité en librairie. Dans le domaine de la littérature, l'offre pourrait se concentrer sur les ouvrages de fonds des éditeurs contemporains et les ouvrages classiques.

UN CHANGEMENT DE PARADIGME

La dynamisation de la fréquentation de la BnF passe par un changement de paradigme dans la prise en compte des besoins et attentes des usagers. Le prêt numérique pourrait contribuer à changer l'image de l'établissement, encore parfois perçu comme une bibliothèque « de dernier recours » ne s'adressant qu'à une population limitée de chercheurs privilégiés, en proposant une offre de qualité sur le plan documentaire, simple et efficace par son ergonomie. En fournissant des données sur les pratiques de ces lecteurs, le prêt numérique permettrait en outre, après anonymisation, de mieux connaître leurs usages, leurs goûts et leurs attentes et, in fine, d'adapter la politique documentaire et le service à leurs besoins.

La transition numérique de la filière de l'édition s'accompagne d'une adaptation de l'offre de prêt en bibliothèque.

03.N L'innovation pour la consultation de la presse

Sur Gallica, la page « Collections » donnant accès aux principaux titres de quotidiens est le plus utilisé de tous les points d'entrée éditorialisés. Cependant, les remontées régulières des attentes des utilisateurs des pages de presse de Gallica amènent à distinguer plusieurs axes d'amélioration.

DÉCRIRE FINEMENT LES CONTENUS

Le projet européen Europeana Newspapers (2012-2014) a offert un précieux terrain d'étude et d'expérimentation (voir aussi axe 6 fiche B). Grâce à la recherche plein-texte avec segmentation des articles, les journaux sont désormais plus finement décrits et bénéficient d'une classification de classes de page et de contenu (publicités, petites annonces, tableaux, illustrations, légendes). Autre type de traitement avancé, la Reconnaissance des Entités Nommées consiste à rechercher et catégoriser des entités telles que personnes, lieux et organisations. D'abord mises en ligne sur un portail dédié par TEL (*The European library*), ces ressources enrichies ont vocation à alimenter à terme une chaîne thématique dédiée du portail Europeana. Avec 18 millions de pages traitées (dont 3 millions provenaient de la BnF), Europeana Newspapers constituait le plus grand projet d'agrégation pour Europeana en nombre de données traitées, de métadonnées converties et d'images incorporées.

AMÉLIORER LES INTERFACES DE RECHERCHE

L'amélioration des paramètres d'interrogation et de navigation textuelle constituerait une autre évolution. Le rapport de recherche développé pour Gallica constitue un pas dans cette direction : ce service mis à disposition des chercheurs facilite le dépouillement d'une recherche précise en texte intégral en assistant l'usager dans la formulation de sa requête, puis en lui fournissant les résultats sous une forme facilement exploitable (rapport de synthèse chiffré, références et lien, extraits imprimables du corpus).

CRÉER DE NOUVELLES FONCTIONNALITÉS

Il serait opportun de s'inspirer des fiches créées pour chaque titre dans data.bnf.fr pour développer Gallica : index nominal de l'ensemble des titres présents, introduction d'une fonction « kiosque », etc.
Il reste à travailler sur les lacunes des collections numérisées (absence dans la collection papier, fichiers numériques manquants ou défectueux), que le développement de l'interopérabilité pourrait contribuer à combler en favorisant les liens vers des sites partenaires dont les collections complètent celles de la BnF, sans que l'usager n'ait à sortir de Gallica.

Europeana Newspapers représente le plus grand projet d'agrégation pour Europeana en nombre de données traitées, de métadonnées converties et d'images incorporées.

Rétronews, un portail dédié

En partenariat avec Immanens, et en collaboration avec Syllabs, BnF-Partenariats (voir axe 1 fiche D) a conçu une offre innovante autour de la presse numérisée : Rétronews. 3,5 millions de pages nouvellement numérisées devraient venir s'ajouter aux 5,5 millions de pages de presse déjà disponibles dans Gallica.
L'offre premium payante apportera des fonctionnalités complémentaires de partage et de téléchargement, des outils de recherche sémantiques et des contenus éditorialisés.

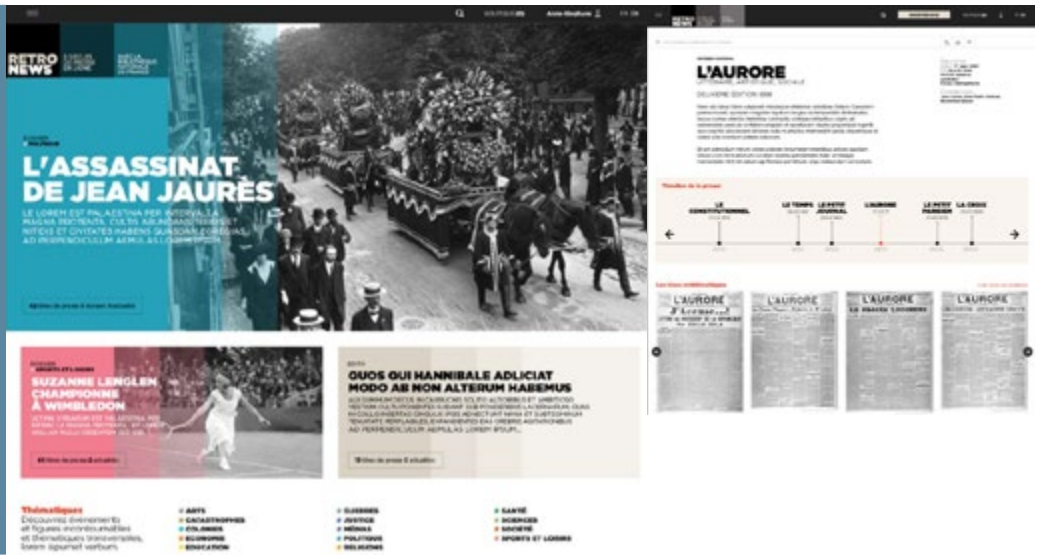


Schéma numérique - mars 2016

03.0 Le déploiement du Wi-Fi

À partir de 2016, les lecteurs de la Bibliothèque pourront, sur l'ensemble de ses sites, accéder à internet via le réseau Wi-Fi, et ce, dans le respect scrupuleux des précautions sanitaires qu'impose l'usage de ces technologies.

FACILITER LA MOBILITÉ

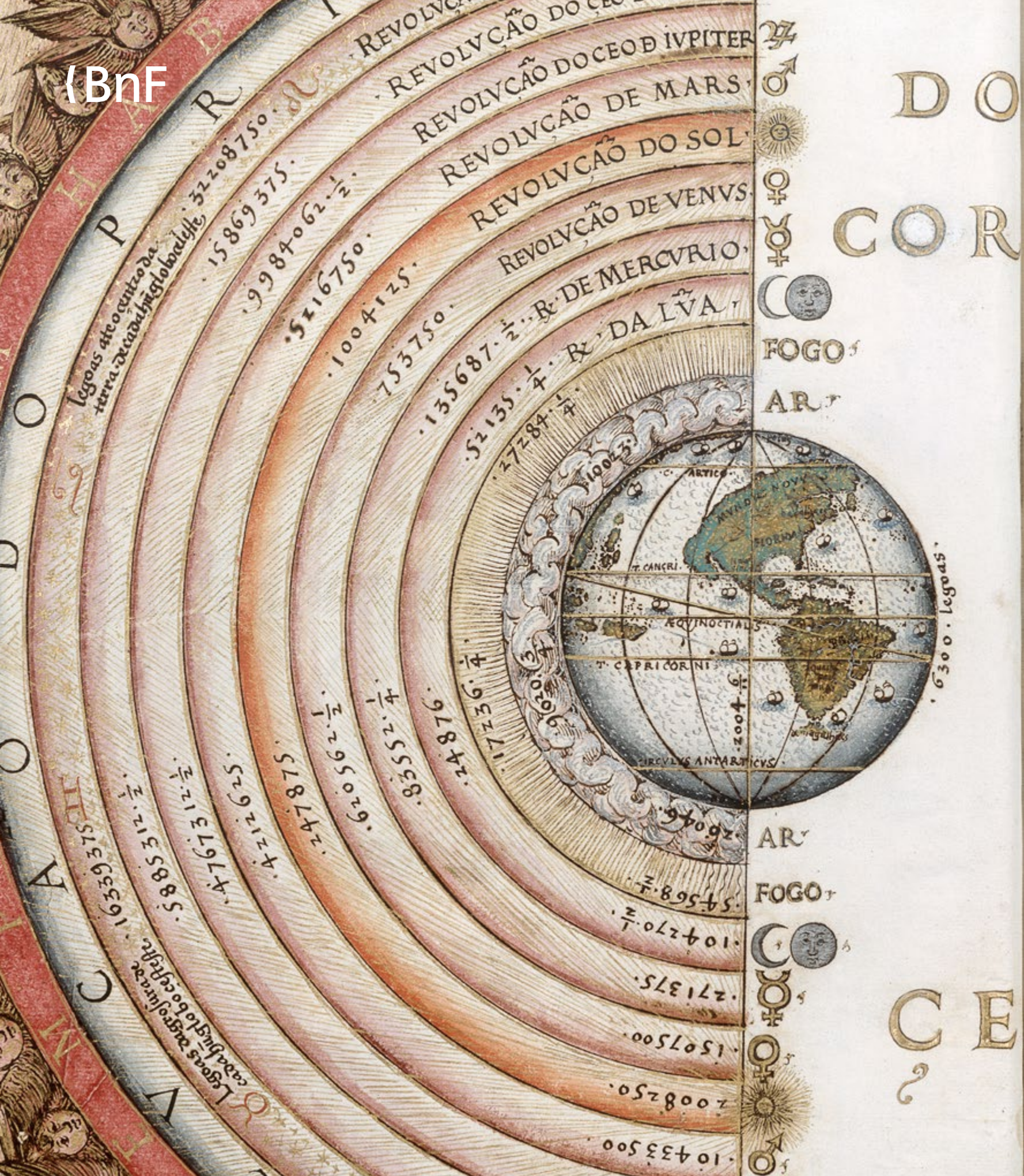
Cette offre, qui fait suite à l'évolution entreprise par la BnF depuis 2008 afin d'offrir à ses lecteurs un accès filaire à internet (avec près de 800 places équipées en 2010 en Bibliothèque de recherche, en Rez-de-jardin (site François-Mitterrand), et près de 500 places équipées fin 2011 en Haut-de-jardin), permettra à partir de 2016 aux lecteurs de la BnF de travailler directement avec leur équipement personnel – que ce soit un PC, une tablette ou un téléphone portable – et de devenir des lecteurs « mobiles ».

Dans le respect des normes sanitaires et afin de minimiser le champ électromagnétique, la puissance maximale d'émission des bornes Wi-Fi est limitée.

… DANS UN SOUCI DE SÉCURITÉ SANITAIRE

Le Wi-Fi sera disponible dans l'ensemble des espaces publics de la BnF (François-Mitterrand, Richelieu, Arsenal, Louvois, Opéra, Avignon), dans les salles de lecture, expositions, halls et auditoriums, à l'exception de zones sans Wi-Fi pour l'accueil des personnes électrohypersensibles.

Dans le respect des normes sanitaires et afin d'avoir un déploiement qui minimise le champ électromagnétique, la puissance maximale d'émission des bornes Wi-Fi est limitée à 30 milliwatts (0 à 30 milliwatts) (soit 30 % de la puissance nominale des antennes à 2,4 gigahertz et 15 % de celle des antennes à 5 gigahertz), et la distance entre les bornes et les usagers, ou entre les bornes et le personnel posté de la BnF, est au minimum de 3 mètres pour une place de lecture, et de 5 mètres pour une banque de salle ou un bureau. Une signalétique claire sera mise en place pour indiquer les zones avec et sans Wi-Fi.



BnF
REVOLVÇÃO DO CEO D IVPITER
REVOLVÇÃO DE MARS
REVOLVÇÃO DO SOL
REVOLVÇÃO DE VENVS
R. DE MERCVRIO
R. DA LVA
FOGO
AR
FOGO
DO
COR
CE
ARTICO
T. CANCRI
AEQVINOCTIALIS
T. CAPRICORNI
CIRCVLVS ANTARTICVS
6300 legoas

04

IRRIGUER UN RÉSEAU DE PARTENAIRES NATIONAUX ET INTERNATIONAUX

145
conventions passées en 2015 avec des partenaires membres du Réseau national de coopération

6 millions
d'euros attribués en soutien financier à 176 partenaires de numérisation entre 2009 et 2014

30 millions
de notices dans le CCFR

+ de 480 000
visites du CCFR en 2015

3 366
fonds documentaires signalés dans le RNBFD

21
projets internationaux de coopération numérique

Grâce au numérique, la Bibliothèque nationale de France fait de son exceptionnel patrimoine un instrument de dialogue entre les territoires et entre les cultures. L'expertise unique qu'elle a su développer dans la production, la collecte, le traitement et la conservation des collections numériques lui permet de jouer un rôle de premier plan dans l'activité scientifique et technique internationale en développant des partenariats avec d'autres établissements culturels à travers le monde, dans l'espace francophone en particulier. Grâce à une politique de coopération nationale dynamique, la BnF contribue à la mise en commun des ressources documentaires des bibliothèques françaises et assure l'animation d'un vaste réseau d'établissements partenaires de programmes de numérisation et de signalement (conversions rétrospectives, catalogage, recensement de fonds, etc.). La BnF mène enfin une politique de recherche ambitieuse qui participe de son rayonnement national, européen et international.

ZOOM SUR QUELQUES POINTS FORTS

04.A_ Les programmes nationaux de numérisation concertée
04.B_ Le catalogue collectif de France (CCFr)
04.C_ L'offre de tiers archivage
04.D_ La coopération numérique internationale
04.E_ Europeana, la bibliothèque numérique du patrimoine culturel européen

MOYENS, OUTILS ET MÉTHODES

04.F_ Gallica marque blanche
04.G_ L'action dans les instances internationales
04.H_ La rubrique « Pour les professionnels » du site institutionnel

AXES DE DÉVELOPPEMENT

04.I_ L'accès en région aux archives de l'internet
04.J_ La diffusion des documents numérisés par d'autres institutions patrimoniales

Un réseau national de partenaires

Afin d'œuvrer à la constitution d'un patrimoine numérique national, la BnF s'est engagée dès 2009 dans une démarche de coopération numérique avec de nombreuses bibliothèques et institutions documentaires publiques (collectivités territoriales, universités, instituts de recherche) et des acteurs privés (fondations, associations, etc.). Cette démarche, prenant appui sur le Schéma numérique des bibliothèques remis en mars 2010, inscrit les bibliothèques françaises dans un projet documentaire numérique collectif à travers des programmes thématiques et d'intérêt régional.

Les enjeux culturels et scientifiques de cette entreprise collective sont pluriels :

- offrir au public un accès aisé au patrimoine national, régional, local ;
- rendre plus visibles sur internet la culture et la langue françaises, en atteignant rapidement une masse critique de données ;
- favoriser la dissémination des ressources numériques et encourager leur appropriation par le public ;
- mettre à disposition des chercheurs des sources méconnues.

La BnF a progressivement élargi le champ de la coopération numérique en proposant aux institutions partenaires de mettre en ligne, dans Gallica, leurs propres ressources numériques patrimoniales. Le développement de l'offre de bibliothèque numérique en Gallica marque blanche répond au besoin de mutualiser les investissements consentis pour le développement de la bibliothèque numérique, en déclinant sous forme de service l'interface de Gallica, aux couleurs des partenaires, et en intégrant leurs propres documents numérisés.

Pivot de la politique de coopération numérique de la Bibliothèque, Gallica a encouragé la formation d'un écosystème de partenaires numériques : ils sont ainsi 305, de statuts, tailles et localisations très variés, à mettre à disposition dans Gallica des centaines de milliers de documents.

La BnF s'est engagée dès 2009 dans une démarche de coopération numérique avec de nombreuses bibliothèques, institutions documentaires publiques et acteurs privés

Acteur central de la mise en œuvre du Plan d'Action pour le Patrimoine Écrit lancé en 2005 par le ministère de la Culture et de la Communication, la BnF poursuit également sa politique de soutien à l'informatisation des catalogues des fonds anciens, locaux et spécialisés, et veille à l'enrichissement continu du Catalogue collectif de France.

réseau BnF
coopération

La BnF au cœur d'un écosystème international

Les collections patrimoniales d'origine étrangère, présentes en grand nombre dans les collections de la BnF, contribuent de manière décisive à son universalisme et constituent un puissant moteur pour les relations internationales. Mais elles lui donnent également une responsabilité toute particulière en matière de description, de conservation physique et numérique, de diffusion et de valorisation de ce patrimoine universel. Elles sont aujourd'hui le socle d'une coopération fondée sur l'idée du partage de ces ressources documentaires exceptionnelles.

Ainsi, depuis 2005, la BnF mène un ambitieux programme de coopération numérique avec les bibliothèques nationales et de recherche étrangères. S'appuyant sur une politique de valorisation numérique des contenus comme mode d'accès à un patrimoine partagé, ces projets favorisent la constitution d'ensembles thématiques donnant accès à des documents dispersés entre les différents pays et institutions partenaires, permettant ainsi de reconstituer un héritage commun. Éclairant d'un autre regard l'histoire et le patrimoine français, ils contribuent à enrichir l'offre documentaire de l'établissement. Si plusieurs de ces programmes ont une dimension essentiellement bilatérale (France-Japon, France-Pologne, France-Brésil, France-Chine, France-Russie), certains couvriront des ensembles géographiques plus larges, ayant un lien historique fort avec la France (bibliothèques du Levant, Indochine). Des mécénats de compétences avec des institutions étrangères permettent, par ailleurs, la description et la numérisation d'ensembles documentaires prestigieux conservés par la BnF : l'International Dunhuang Project, qui réunit plusieurs collections étrangères issues du site de Dunhuang (Chine) ; le projet de numérisation et d'édition en ligne des imprimés chinois anciens conservés dans les grandes bibliothèques étrangères mené par l'Université du Shandong (Chine) ; le projet de numérisation de la presse ottomane en langues allophones piloté par l'Institut français d'études anatoliennes et l'institution turque SALT Research.

La BnF exerce une responsabilité toute particulière en matière de description, de conservation physique et numérique, de diffusion et de valorisation d'un patrimoine universel.

Enfin, c'est aussi au niveau international que se jouent les évolutions majeures des bibliothèques dans le cadre d'une transition numérique dont les enjeux sont globaux. L'expertise de la BnF lui vaut une place de premier plan dans ces instances où convergent les dynamiques internationales de l'innovation, comme l'IFLA (*International Federation of Library Associations*), le consortium IIPC (*International Internet Preservation Consortium*) ou la fondation Europeana, entre autres.

Les relations avec le monde de la recherche

Les domaines d'excellence de la BnF au regard du numérique – préservation, numérisation, bibliothèque numérique, archivage du web, web sémantique, médiation numérique, etc. – ont vocation à servir un spectre élargi d'acteurs professionnels, qui sont autant d'ambassadeurs potentiels des savoir-faire de l'établissement : étudiants, chercheurs et professionnels de la documentation et de l'information ou, plus largement, tous les professionnels de la recherche et du développement dans les domaines des services et de l'industrie.

La BnF s'est constituée en acteur privilégié dans un paysage de la recherche nationale en pleine restructuration.

À la faveur de son partenariat avec cinq laboratoires d'excellence (LabEx Arts-H2H, PATRIMA, CAP, PP, OBVIL), trois équipements d'excellence (ÉquipEx Biblissima, Patrimex et Ortolang) et le CNRS, la BnF s'est constituée en acteur privilégié dans un paysage de la recherche nationale en pleine restructuration. Cette position lui permet de mener des collaborations scientifiques et d'explorer des champs d'étude originaux, qui sont autant d'opportunités de valoriser les importantes collections numérisées ces dernières années.

Plusieurs projets de recherche et développement sont ainsi initiés en lien avec des équipes pluridisciplinaires. Un modèle de coopération numérique avec le monde universitaire et le monde de la recherche se dessine autour des expertises de la BnF en matière de nouveaux usages (pratiques de lecture numérique, réseaux sociaux, production participative (*crowdsourcing*)…) et de mégadonnées (*big data*). Voués à créer les conditions d'une nouvelle culture partagée, les humanités numériques, ces programmes traitent notamment de problématiques d'indexation sémantique de données, de cartographie des informations, d'indexation de séquences audiovisuelles, d'indexation automatique sélective ou d'indexation assistée par ordinateur (IAO).

MER DE L'INDE
L'EQUATEUR
CLAUDE DU TROPIQUE DE CAPRICORNE
HET NIEW HOLLANDT
NOUVELLE HOLLANDE
MERIDIONAL
INCONNUE
ANTARCTIQUE

04.A Les programmes nationaux de numérisation concertée

La numérisation partenariale s'élabore selon une politique documentaire guidée par des principes structurants : numérisation d'imprimés en français et dans les langues de France entrés dans le domaine public, ou dont les droits ont été négociés, principalement dans le cadre de programmes nationaux de numérisation. Le but poursuivi est la constitution de corpus thématiques ou d'intérêt régional, destinés à être très visibles sur internet.

UNE LOGIQUE DE CORPUS

Les bibliothèques françaises sont invitées à participer à plusieurs programmes de numérisation concertée autour d'ensembles thématiques (sciences juridiques, histoire de l'art, guerre de 1914-1918, littérature patrimoniale pour la jeunesse), de fonds d'intérêt régional (histoire des territoires, sociétés savantes locales et régionales) ou de corpus plus ciblés, précieux sur le plan documentaire (publications des académies et sociétés nationales, journaux de tranchées, presse clandestine de la Résistance, presse de l'immigration). La Bibliothèque conduit également des programmes de numérisation concertée autour de ressources non imprimées (portulans, manuscrits), avec l'objectif de reconstituer virtuellement des fonds dispersés.

LES DISPOSITIFS DE SOUTIEN FINANCIER

Pour donner de l'envergure à cette entreprise partenariale et garantir son succès, la BnF a mis en place un dispositif de soutien financier incitatif : attribution de subventions pour les projets de numérisation ou intégration d'imprimés patrimoniaux des bibliothèques françaises dans les marchés de numérisation de la BnF, ouverts pour un tiers aux collections d'autres bibliothèques.

L'efficacité de l'action conduite par la BnF se mesure à l'aune des résultats obtenus : entre 2009 et 2014, 17,38 millions de pages d'imprimés ont été numérisées par 176 partenaires, soutenus financièrement à hauteur de 6,1 millions d'euros ; une action appelée à se poursuivre et à s'amplifier sur la période 2015-2019. L'évolution du contexte numérique national et de l'action des ministères de la Culture et de la Communication, d'une part, et de l'Enseignement supérieur et de la Recherche, d'autre part, appelle néanmoins la recherche d'une articulation optimale des dispositifs de financement.

L'aide à la numérisation s'accompagne d'exigences renforcées en matière d'accès aux ressources numérisées (OCR, structuration, indexation) et de mise en valeur partagée. Une attention particulière est accordée à l'optimisation de la diffusion des documents numériques, de manière à donner la plus grande visibilité possible aux ressources.

L'aide à la numérisation s'accompagne d'exigences renforcées en matière d'accès aux ressources numérisées et de mise en valeur partagée.

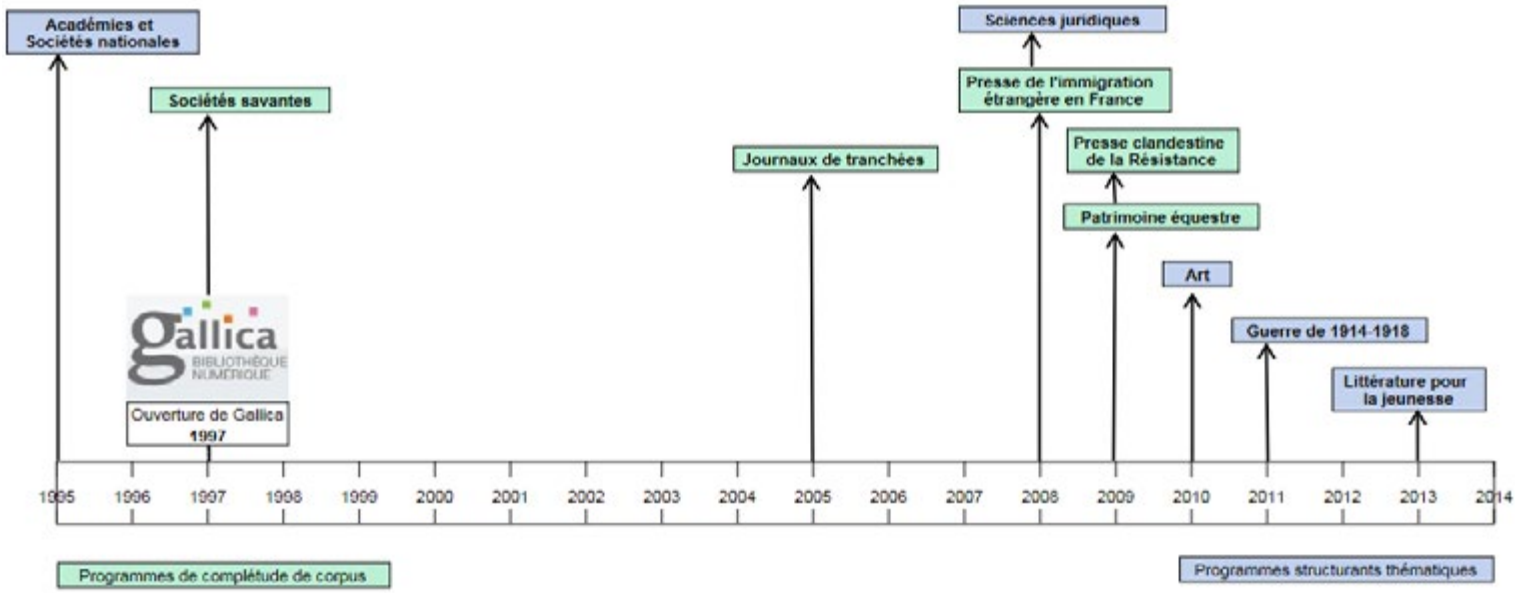


Schéma numérique - mars 2016

04.B
Le catalogue collectif de France (CCFr)

Dans le cadre de sa politique de coopération, la BnF produit le Catalogue collectif de France (CCFr) qui donne accès à 30 millions de références conservées dans plus de 3 500 bibliothèques. Fédérant les principaux catalogues français, il offre une visibilité élargie aux ressources de ces bibliothèques et constitue un précieux outil de travail pour les chercheurs et professionnels du livre.

TROIS ENSEMBLES DOCUMENTAIRES D'UNE GRANDE VALEUR SCIENTIFIQUE

Trois ensembles composent l'offre du CCFr :

- **la Base patrimoine** décrit et localise 5,7 millions de documents patrimoniaux conservés dans 182 bibliothèques françaises. Cet ensemble est en accroissement permanent sous l'effet des opérations de conversion rétrospective soutenues financièrement par la BnF, ou de versements volontaires.
- **le Répertoire national des bibliothèques et des fonds documentaires (RNBFD)** propose des informations relatives à plus de 5 000 bibliothèques et plus de 3 300 fonds, avec des liens vers les bibliothèques numériques.
- Enfin, le CCFr fait office de **catalogue collectif national des manuscrits** grâce à un accès fédéré au Catalogue général des manuscrits des bibliothèques publiques de France (CGM), à BnF archives et manuscrits et à Calames, le catalogue des manuscrits et archives des bibliothèques de l'Enseignement supérieur et de la recherche.

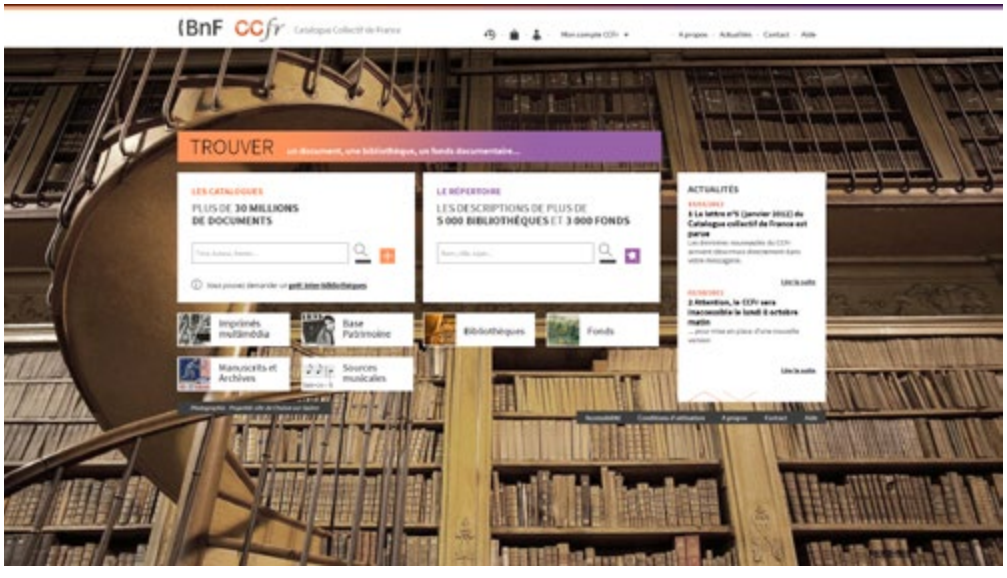
Le CCFr accroît continument le nombre et la variété des données accessibles via son portail ccfr.bnf.fr

UN CATALOGUE EN CONSTANTE ÉVOLUTION

Le CCFr accroît continûment le nombre et la variété des données accessibles via son portail ccfr.bnf.fr. Il développe des outils partagés, permettant à ses partenaires de participer à l'augmentation et à l'enrichissement des ressources (mise à jour participative des notices de fonds et d'institutions dans le RNBFD, outil national de catalogage en EAD en vue d'une actualisation du CGM), mais aussi de réutiliser les données disponibles (export vers des outils de gestion de bibliographie et récupération des notices, permaliens, résolveurs de liens, etc.).
Une autre évolution notable du CCFr consiste à multiplier les liens vers des documents numériques. Centrée dans un premier temps sur les bases de manuscrits, cette démarche permettra d'accéder, depuis les catalogues, à un volume croissant de ressources numériques. À cet égard, le moissonnage par les moteurs de recherche de l'intégralité des données hébergées ou encore l'intégration d'ensembles de données dans des bases spécialisées, comme la base Heritage Printed Books du *Consortium of European research libraries* (CERL), seront dans les prochains axes de travail.

ACCROÎTRE LA VISIBILITÉ DU CCFR

Il est essentiel que la présence du CCFr sur le web soit accentuée afin d'offrir aux institutions qu'il fédère une visibilité croissante. Par son ergonomie et ses fonctionnalités renouvelées, la nouvelle version du site (2016) devrait sensiblement y concourir, accélérant par ailleurs la dissémination des données.



Projet de nouvelle interface du CCFr (2016)

04.C L'offre de tiers archivage

Dès la création de SPAR (voir axe 2 fiche B) s'est posée la question de l'ouverture du système à d'autres organisations confrontées aux mêmes enjeux de conservation et souhaitant bénéficier des technologies et du savoir-faire de la BnF. Dans ce domaine particulièrement évolutif, qui nécessite le renouvellement régulier des technologies, la mutualisation est une solution pour les organisations qui ne peuvent supporter les coûts de façon isolée.

BNF ARCHIVAGE NUMÉRIQUE

Pour répondre à ces enjeux, la BnF propose aujourd'hui l'offre BnF archivage numérique, un service de tiers archivage offrant à un tarif juste et modéré les mêmes garanties de sécurité et de pérennité que celles en vigueur pour ses propres collections patrimoniales.

DES INITIATIVES STRUCTURANTES

De manière à offrir un niveau et une qualité de service comparables à ceux du secteur privé, la BnF a pris plusieurs initiatives. L'agrément du SIAF (Service interministériel des Archives de France) a été obtenu en 2013, permettant à la BnF de se positionner comme opérateur de confiance. L'implémentation de MEDONA (Modélisation des échanges de données pour l'archivage), devenue norme NF Z44-022, a été entreprise. Elle vise à normaliser les messages entre gestionnaires de flux d'archives électroniques (services versants, services d'archives…). La Bibliothèque a enfin fait évoluer l'extranet Espace Coopération (voir axe 5 fiche C) afin de faciliter les opérations de versement et l'accès aux données pour les bénéficiaires du tiers archivage.

Le Centre Pompidou archive ses données dans SPAR depuis 2013 (10 To archivés en 2015). Le musée Picasso et le musée du Quai Branly ont également choisi la BnF pour assurer l'archivage de leurs collections numériques (premiers versements en 2015).

Tout en se positionnant comme opérateur de confiance, la BnF facilite les opérations de versement et l'accès aux données.

Les tarifs

Ils sont établis en fonction de la volumétrie de stockage, du mode de téléchargement des données à archiver, du nombre de copies et du type de supports de stockage, de la durée du contrat.

Un archivage sur bandes offre un support peu coûteux, mais lent. Sur disques, la disponibilité des données sera plus rapide.

Les recettes se montent à 40 000 euros en 2015.

Exemple du tarif pour un contrat de cinq ans, avec un archivage minimal sur deux bandes :

Bnf Archivage numérique
Tarif général annuel, au téraoctet, en € HT, basé sur 5 ans

Volume des données à archiver sur deux bandes	extranet	sur mesure
pour 1 To	2 545 €	3 990 €
de 2 à 5 To	1 960 €	2 680 €
de 6 à 9 To	1 740 €	2 160 €
de 10 à 29 To	1 590 €	1 840 €
de 30 à 49 To	1 420 €	1 530 €

ARCHIVAGE NUMÉRIQUE

BnF Bibliothèque nationale de France

04.D
La coopération numérique internationale

COUVRIR DE NOUVEAUX ENSEMBLES GÉOGRAPHIQUES

Les programmes numériques bilatéraux incitent à constituer des ensembles thématiques à partir de documents dispersés dans différents pays et institutions. Ils contribuent ainsi à reconstituer une histoire et des patrimoines liés.

Depuis 2005, la BnF a initié de nombreux programmes de coopération numérique, dans le cadre de tels partenariats : la France en Amérique, France-Brésil / A França no Brasil, France-Japon, « Sinica », France-Pologne, Presse ottomane en français…

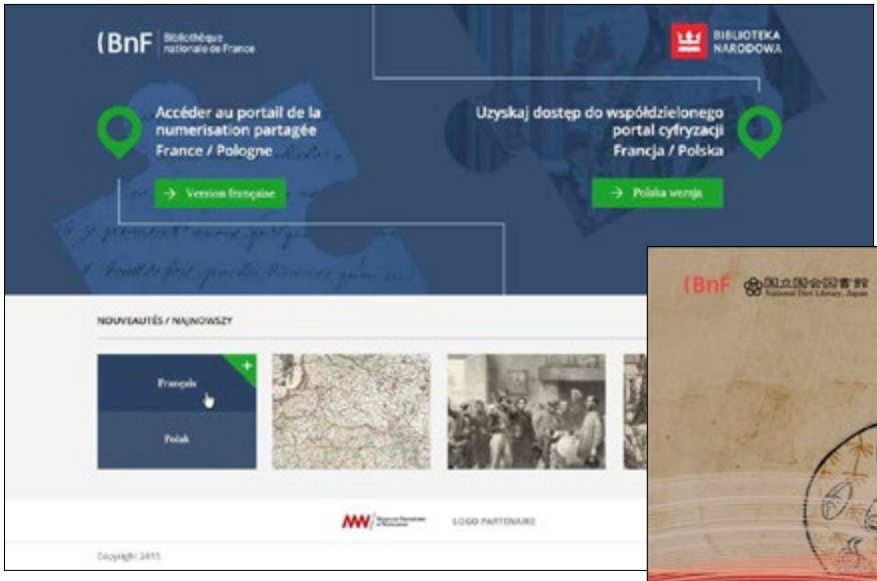
La BnF contribue aussi à des programmes collaboratifs multilatéraux (World Digital Library, International Dunhuang Project) ainsi qu'à des programmes multilatéraux menés par des institutions nationales (livres anciens chinois, Taïwan). Elle est régulièrement sollicitée pour participer à de nouveaux projets (université de Shandong, manuscrits indiens, bibliothèque numérique qatarie).

Certains de ces programmes ont vocation à couvrir des ensembles géographiques ayant un lien historique fort avec la France comme en témoignent le soutien apporté à la création d'une bibliothèque numérique des bibliothèques françaises du Levant (école biblique de Jérusalem, Dominicains du Caire, bibliothèque Saint-Joseph au Liban…) ou le développement d'un portail consacré à l'Indochine française.

Les contenus produits par ces programmes numériques sont pérennisés en étant intégrés à Gallica, dont l'infrastructure technologique permet de mutualiser progressivement les briques fonctionnelles (éditorialisation sous Drupal, Gallica marque blanche).

CONTRIBUER AU RAYONNEMENT DU FRANÇAIS

Ces parcours thématiques enrichissent l'offre patrimoniale de la BnF en direction des chercheurs et des amateurs français et étrangers, participant pleinement de la démarche de médiation numérique engagée par l'établissement. Ils sont aussi les relais d'une politique résolue en faveur du français comme langue de culture.



Exemples de portails bilatéraux de la collection « Patrimoines partagés » : France-Pologne, France-Japon

04.E Europeana, la bibliothèque numérique du patrimoine culturel européen

Créée en novembre 2008 avec le soutien de la Commission européenne, Europeana agrège les métadonnées de 33 millions d'objets numériques (textes, images, sons, vidéos) provenant de 2 000 institutions européennes (bibliothèques, archives, musées…).

UN RÔLE MAJEUR POUR LA BNF

Son statut de membre fondateur et l'élection en 2011 de Bruno Racine, président de la BnF, à la tête de la Fondation Europeana, renouvelée en 2013 et 2015, permettent à la BnF de jouer un rôle majeur dans la définition de la nouvelle stratégie 2015-2020 : s'imposer comme vitrine du patrimoine culturel européen tout en faisant évoluer le portail vers une plateforme de distribution de données à l'usage des industries créatives, des fournisseurs de contenus ou tout simplement des citoyens.

LA VITRINE D'UN SAVOIR-FAIRE

La BnF accompagne fortement cette évolution en apportant à Europeana son expertise en matière de gestion des données et en faisant bénéficier le consortium des avancées technologiques développées dans le cadre de data.bnf.fr. Elle a longtemps été le premier fournisseur de contenus d'Europeana, grâce au partage des métadonnées de l'ensemble des documents disponibles dans Gallica. Elle est également l'agrégateur vers Europeana des bibliothèques françaises partenaires de Gallica, favorisant ainsi la dissémination au niveau européen des données produites sur l'ensemble du territoire français.

PRÉSENTER LES GRANDES ŒUVRES DE LA CULTURE EUROPÉENNE

La BnF contribue fortement à l'enrichissement de l'offre de médiation culturelle du portail, à travers Europeana 1914-1918, dédié à la Première Guerre mondiale, ou encore le projet Europeana Sounds, consacré au patrimoine sonore européen.
La constitution d'un ensemble éditorialisé présentant les grandes œuvres de la culture européenne devrait pour sa part s'appuyer sur l'expertise développée par la Bibliothèque avec les Essentiels de la littérature (voir axe 3 fiche B).

La BnF joue un rôle d'agrégateur des bibliothèques françaises vers Europeana.

europeana
think culture

europeana
newspapers

europeana
sounds

04.F
Gallica marque blanche

Le dispositif Gallica marque blanche est né d'une double volonté : mutualiser les développements réalisés dans le cadre de Gallica pour en faire bénéficier les partenaires de la Bibliothèque ; compléter les collections nationales en intégrant des documents numérisés par les autres bibliothèques françaises.

UNE BIBLIOTHÈQUE NUMÉRIQUE PERSONNALISÉE

Chaque partenaire de l'offre Gallica marque blanche dispose de sa propre bibliothèque numérique, tandis que la BnF améliore le service rendu sur Gallica en proposant un fonds documentaire enrichi, également accessible dans Gallica intra muros. Tout document présent dans une bibliothèque Gallica marque blanche dispose d'une notice dans les catalogues de la BnF. Hormis la page d'accueil, spécifique à chaque site, le processus de recherche et de consultation des documents est identique et s'appuie sur les mêmes outils et les mêmes technologies. Les pages sont personnalisées aux couleurs du partenaire, qui appose son logo et garde la possibilité de paramétrer l'activation ou la désactivation des différentes fonctionnalités proposées. Par ailleurs, la BnF assure la maintenance et l'hébergement de chaque site en marque blanche, même si celui-ci dispose de sa propre URL, indépendante du domaine bnf.fr.

UNE DÉMARCHE DE RATIONALISATION

La BnF s'est engagée en 2015 dans une démarche de rationalisation et de développement du dispositif Gallica marque blanche, pour en faire bénéficier le plus largement possible les établissements intéressés. À la suite de cette étape et de la nouvelle infrastructure technique et fonctionnelle liée à la refonte de Gallica, de nouveaux projets peuvent voir le jour en 2016 dans un contexte simplifié.

DES BÉNÉFICES PARTAGÉS

Ce dispositif constitue pour la BnF une formidable opportunité de développement de la collection numérique, tout en permettant à d'autres acteurs publics de partager le bénéfice de plusieurs années d'investissements informatiques et d'expertise technologique. Le développement de futurs projets en marque blanche privilégiera des partenaires dont les fonds présentent un intérêt documentaire majeur et une infrastructure technique adéquate, avec une attention particulière accordée à leur diversité.

Un modèle économique a été défini afin d'établir un équilibre entre les investissements consentis par la BnF et la contribution des partenaires.

Deux sites en marque blanche : Numistral (Bibliothèque nationale universitaire de Strasbourg, 2013) et La Grande Collecte (partenariat avec le Service interministériel des Archives de France, 2014)

04.G L'action dans les instances internationales

La BnF est fortement impliquée dans les grandes instances patrimoniales internationales, contribuant de manière significative à l'innovation et au développement des bibliothèques, dans le champ numérique notamment.

UNE PRÉSENCE INTERNATIONALE ACTIVE

Assurant depuis sa fondation (2003) un rôle majeur dans la gouvernance du Consortium international pour la préservation de l'internet (IIPC), la BnF participe très activement à celle du Fichier d'autorité international virtuel (VIAF).
Très présente dans la Fédération internationale des associations de bibliothécaires et des bibliothèques (IFLA), comme dans la Ligue des bibliothèques européennes de recherche (LIBER), son expertise est particulièrement attendue sur les questions de dépôt légal du web, de normalisation bibliographique et d'échanges de données, ou encore de conservation pérenne. La BnF participe au comité technique d'IASA (*International Association of Sound and Audiovisual Archives*), où sont rédigées les préconisations de format de conservation des documents audiovisuels. Dans le cadre de la présidence de la Conférence européenne des directeurs de bibliothèques nationales (CENL), la BnF a largement œuvré à la définition d'une nouvelle offre de services destinée aux bibliothèques de recherche (traitement et fouille de données).

CONTRIBUER AUX TRAVAUX DE NORMALISATION

La BnF collabore enfin à des travaux de normalisation décisifs, notamment à l'ISO, où elle a piloté la normalisation du format de stockage et de préservation des archives du web, WARC (ISO 28500 : 2009), ainsi qu'un rapport technique sur les statistiques et indicateurs de qualité des archives du web (ISO/TR 14873).

L' expertise de la BnF est particulièrement attendue sur les questions de dépôt légal du web, de normalisation bibliographique, d'échanges de données ou encore de conservation pérenne.

Le Réseau francophone numérique

RFN est un consortium regroupant vingt-cinq institutions documentaires de la Francophonie.
Sa mission est d'assurer la préservation et l'accessibilité des contenus numériques francophones.
La coopération numérique est l'un de ses principaux outils, avec l'expertise et la formation. Pour atteindre ses objectifs, le consortium a élaboré un nouveau plan stratégique, autour de deux ambitions majeures : soutenir et accélérer les programmes de numérisation et la constitution de bibliothèques numériques propres ; faire évoluer le portail, avec une offre de médiation à forte valeur ajoutée.
La BnF contribue significativement en supervisant l'évolution technique et l'enrichissement du portail.
Elle s'est par ailleurs engagée à développer, à partir de 2016, la bibliothèque numérique francophone collective du RFN via le dispositif Gallica marque blanche.

04.H La rubrique « Pour les professionnels » du site institutionnel

En ligne depuis 2002, la rubrique « Pour les professionnels » est accessible à tous depuis le site institutionnel www.bnf.fr. C'est grâce à ce relais que l'établissement communique auprès de ses publics professionnels et joue ainsi son rôle de tête de réseau national et international. Cette rubrique s'adresse aux bibliothécaires et aux professionnels du livre et de la conservation, mais aussi à d'autres communautés comme celles des archives, du web et des industries numériques. Elle est consultée par un public d'experts comme un outil de veille, mais également par les étudiants et les praticiens dans le cadre de la formation continue, ou pour répondre à des besoins concrets.

Dans le contexte actuel de mutualisation des compétences, de l'*open data* et des changements induits par le numérique, cette rubrique constitue un gisement de ressources utile pour mieux comprendre la BnF « de l'intérieur » et pour accompagner les professionnels vers de nouvelles pratiques de leurs métiers.

Fortement concernée par l'impact du numérique sur l'évolution de ses missions et de ses métiers, la BnF valorise tout particulièrement son savoir-faire dans ce secteur grâce à la rubrique « Innovation numérique » du site www.bnf.fr.

DES INFORMATIONS QUI FONT RÉFÉRENCE

La BnF y partage son expertise dans tous les métiers de la bibliothéconomie. La rubrique « Catalogage et indexation » fait référence en ce qui concerne les normes de catalogage, les formats et leurs évolutions. La rubrique « Conservation » contient de précieuses informations sur la conduite d'un plan d'urgence pour sauvegarder des collections en cas de sinistre.
Les pages « Coopération nationale » informent sur la politique et les programmes de coopération de la BnF.

FOCUS SUR « INNOVATION NUMÉRIQUE »

Fortement concernée par l'impact du numérique sur l'évolution de ses missions et de ses métiers, la BnF valorise tout particulièrement son savoir-faire dans ce secteur grâce à la rubrique « Innovation numérique ».
Les contenus de chacune des thématiques de cette rubrique (dépôt légal du numérique, Gallica, numérisation, préservation numérique, records management et web de données) sont organisés suivant le même modèle : une présentation générale des enjeux, une « boîte à outils » sur les référentiels et les formats utilisés par la BnF, et une partie « Coopération et mise en oeuvre » dans laquelle sont notamment présentés les projets de R&D où la BnF est impliquée, ainsi que des applications phares telles que l'entrepôt numérique SPAR ou le service data.bnf.fr.

DE NOMBREUX SERVICES

Les pages professionnelles fournissent également des informations pratiques sur des services que la BnF rend dans le cadre de missions plus anciennes, comme le dépôt légal ou la récupération des données. Placées sous la « Licence ouverte » de l'État depuis 2014 (voir axe 2 fiche A), les données bibliographiques et les données d'autorité de la BnF sont désormais librement réutilisables.
Pour répondre aux besoins croissants des utilisateurs, la rubrique « Récupération et utilisation des données de la BnF » propose ainsi des fiches pratiques et des informations techniques sur les modalités de récupération.

L'ÉCHO DE RENCONTRES PROFESSIONNELLES PASSÉES OU À VENIR

Enfin, cette rubrique valorise l'actualité professionnelle de la BnF à travers des brèves et un agenda. Les actes, communications et captations vidéo des principales journées d'étude organisées par la BnF y sont accessibles par thématique et par année. On y trouvera notamment tous les actes des journées des Pôles associés et de la coopération ainsi que les captations de la journée annuelle co-organisée avec l'AFNOR sur la normalisation documentaire.

Littérature jeunesse : un nouveau site pour les professionnels

En 2016, les actions, l'offre de formation et les ressources du CNLJ à destination des médiateurs du livre, chercheurs et professionnels en littérature de jeunesse trouveront un nouvel écrin (cnlj.bnf.fr), qui continuera à apporter un éclairage sur l'actualité générale de ce domaine et proposera de très nombreux contenus en ligne : revues spécialisées (*La Revue des livres pour enfants, Takam Tikou*), bibliographies… Ces ressources se veulent aussi accessibles au grand public.

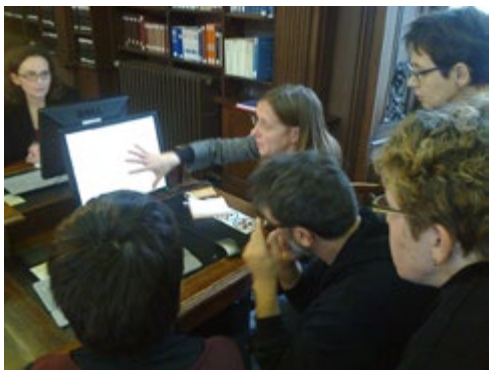
04.1 L'accès en région aux archives de l'internet

La question de la collecte des archives du web a pour corollaire celle de leur visibilité. Il importe de mieux faire connaître les collections ainsi constituées, en leur assurant un signalement en ligne et en mettant les collectes ciblées en valeur sur data.bnf.fr.

RENFORCER LE MAILLAGE

Les points d'accès aux collections sont multipliés au bénéfice des bibliothèques de dépôt légal imprimeur, dont le réseau assure un bon maillage du territoire national. L'arrêté de 2014 fixant la liste des organismes habilités à mettre en œuvre la consultation sur place des services de communication au public en ligne collectés au titre du dépôt légal autorise la BnF à déployer cet accès. Un programme de formation spécifique ainsi que des supports de communication adéquats sont proposés aux partenaires. Les bibliothèques de Montpellier, Nancy, Strasbourg, Lyon, Limoges, Toulouse et Bordeaux en ont bénéficié en 2014-2015.

Dans chaque bibliothèque de dépôt légal imprimeur, l'accès aux archives de l'internet se fait selon les mêmes modalités qu'à la BnF : il est ainsi réservé aux chercheurs accrédités.



Formation à la consultation des archives de l'internet à la médiathèque de Nancy

SÉCURISER L'ACCÈS AUX DONNÉES

Conformément aux dispositions légales, chaque bibliothèque de dépôt légal imprimeur doit réserver l'accès aux archives de l'internet aux chercheurs accrédités.
Pour garantir la sécurité de l'accès à ces collections, la BnF utilise une solution de virtualisation : l'accès s'effectue directement à distance sur les serveurs de la BnF, sans qu'il soit nécessaire de créer une copie locale des données.

LA COLLECTE DE L'INTERNET LOCAL

Cet accès vient compléter une politique de coopération active avec les régions sur le thème de l'archivage de l'internet (voir axe 1 fiche A) : depuis 2004, une sélection partagée des sites web destinés à la collecte a lieu à l'occasion des campagnes électorales (21 bibliothèques participantes pour les élections régionales 2015). La bibliothèque nationale et universitaire de Strasbourg a été la première à s'associer de manière pérenne à la constitution d'une collecte du web d'intérêt local et régional : les « Alsatiques du web ». Avec l'accès distant à ces archives, c'est donc un patrimoine local de premier plan, comprenant également la presse quotidienne régionale collectée sous forme numérique, qui devient disponible en région pour les lecteurs accrédités.

04.J La diffusion des documents numérisés par d'autres institutions patrimoniales

L'enrichissement des fonds accessibles grâce à Gallica s'opère suivant différents processus. Certains, partenariaux, permettent de signaler et de diffuser les ressources numériques créées par d'autres institutions patrimoniales.

DEUX MODES D'ENTRÉE

Gallica prévoit deux modes d'entrée de ressources numériques externes :

- **le référencement par interopérabilité** (protocole OAI-PMH) : en novembre 2015, plus de 70 bibliothèques numériques sont signalées ainsi dans Gallica ;
- **l'intégration par dépôt de fichiers numériques** : cette solution, expérimentée depuis 2010, vaut pour les partenaires ne disposant pas de bibliothèque numérique. Exigeant à l'origine une entrée des fichiers à l'unité, le système a été soumis à une expérience d'intégration de masse, dans le cadre du projet Numistral conduit avec la bibliothèque nationale et universitaire de Strasbourg.

Si l'interopérabilité s'est avérée particulièrement facile à mettre en oeuvre, l'intégration dans Gallica de fichiers numérisés par les bibliothèques françaises demeure un processus plus complexe, les critères techniques de prise en charge de ces fichiers étant sélectifs afin d'assurer les meilleures conditions de diffusion dans Gallica et de conservation dans SPAR. Pour que les fichiers puissent être intégrés en masse, l'exactitude des métadonnées, leur harmonisation et leur degré de précision doivent faire l'objet d'une attention particulière.

Près de 9 millions de pages d'imprimés issus d'une centaine de bibliothèques partenaires ont été numérisées par la BnF et sont actuellement diffusées dans Gallica.

DIFFUSER ET PÉRENNISER UN PATRIMOINE INEXPLOITÉ

L'intégration de fichiers dans Gallica présente un fort potentiel, car elle permettra d'apporter aux partenaires une solution à un problème identifié depuis plusieurs années : la conservation et la diffusion de leurs collections numériques. Si les dispositifs d'aide au financement de la numérisation ont en effet permis à un certain nombre d'établissements de reproduire leurs fonds, ils ne disposent pas tous pour autant d'une bibliothèque numérique ou d'une interface de diffusion, et nombre de ces fonds constituent un « patrimoine dormant » qu'il convient enfin de diffuser et de valoriser.

Ces contenus de valeur sont présentés sur des pages dédiées aux partenaires dans Gallica. Ils peuvent être consultés et exploités à l'aide de toutes les fonctionnalités de la bibliothèque numérique, à l'instar des quelque 9 millions de pages d'imprimés issus d'une centaine de bibliothèques partenaires qui ont été numérisés par la BnF et sont actuellement diffusés dans Gallica.

Place réservée
VAIO

05

SIMPLIFIER LES OUTILS ET LES PROCESSUS POUR LA BNF ET SES PARTENAIRES

Une progression exponentielle

20 000
documents numérisés en 1998

350 000
documents dans Gallica en 2007

3,5 millions
de documents dans Gallica en 2015

Entre 2005 et 2015,
passage de
16 To
à plus de
668 To
d'archives du web

L'essor du numérique dans la première décennie du XXIe siècle s'est accompagné, à la BnF, d'un véritable changement d'échelle. Avec la numérisation de masse, le nombre de documents numérisés a été multiplié par dix entre 2007 et 2015.

Dans le même temps, les archives de l'internet, qui représentaient seulement 16 téraoctets de données en 2005, ont également connu une forte progression (180 téraoctets fin 2009 et 668 téraoctets fin 2015).

Des croissances exponentielles de même nature sont attendues pour les documents nativement numériques appelés à entrer par dépôt légal, qu'on pense aux milliers de livres numériques attendus prochainement ou aux 178 000 albums de musique désormais majoritairement commercialisés sous une forme dématérialisée.

Cette augmentation considérable, sous l'effet des technologies numériques, des documents produits et collectés a eu plusieurs implications, dont la nécessité d'instituer un système adapté à la réception en nombre des documents, de se doter de moyens de suivi automatisés des flux de documents physiques et numériques, et de redéfinir les processus de contrôle qualité. Il devenait également vital de répondre ●●●

●●● à la demande des partenaires, qu'ils soient institutionnels ou issus du monde marchand, qui attendent de la BnF fluidité et réactivité pour absorber leurs flux, connecter leurs propres outils de production ou de suivi à ceux de l'établissement, et faciliter les interactions automatisées entre les systèmes. De plus, l'adaptation de l'organisation du travail aux changements introduits par le numérique interroge le cadre juridique, tant au plan de la collecte (prise en compte du livre numérique) que des traitements (fouille de données et de métadonnées) ou de l'accessibilité (consultation sur place ou à distance, dans un environnement sécurisé).

Face à ces enjeux qui, en outre, se font jour dans un contexte caractérisé par des moyens humains et financiers contraints, la BnF s'est dotée d'outils visant à adapter et simplifier ses processus de travail interne comme ses échanges avec des tiers (éditeurs, bibliothèques, prestataires).

ZOOM SUR QUELQUES POINTS FORTS

05.A_ La chaîne de numérisation : processus et système
05.B_ L'extranet du dépôt légal
05.C_ L'extranet Espace Coopération
05.D_ La plateforme PLATON pour l'exception handicape

MOYENS, OUTILS ET MÉTHODES

05.E_ La gestion électronique des documents
05.F_ Les outils de l'archivage du web
05.G_ La qualité

AXES DE DÉVELOPPEMENT

05.H_ Les outils de gestion des collections numériques
05.I_ Les acquisitions et dons de collections numériques (ADDN)

La dématérialisation des procédures administratives

La Bibliothèque nationale de France a anticipé dès le milieu des années 1980 l'impact du numérique sur la majorité des processus bibliothéconomiques, à commencer par le catalogage, qu'elle a informatisé. Quand la nécessité de passer d'une gestion individuelle des documents administratifs internes à une organisation plus collective s'est imposée, au milieu des années 2000, et que les projets transverses associant de multiples compétences internes se sont multipliés, la BnF a engagé la dématérialisation des procédures jugées mûres : appel d'offres, relations avec certains prestataires, gestion des identités et équipements du personnel, chaîne des recettes et des dépenses… La BnF est ainsi le premier établissement public administratif à gérer exclusivement des documents dématérialisés au sein de sa chaîne de gestion financière, ce qui génère une économie annuelle d'une tonne et demie de papier et un raccourcissement des délais de paiement de 6 jours en moyenne par dossier.

La BnF est le 1er établissement public administratif à gérer exclusivement des documents dématérialisés au sein de sa chaîne de gestion financière.

Ces projets ont cherché à concilier au mieux les intérêts stratégiques de l'établissement et le bon fonctionnement des activités quotidiennes individuelles : en facilitant la création et la transmission des documents ainsi que l'ensemble des étapes de validation ; en assurant la validité administrative et juridique des documents et des données, afin de préserver les droits des personnes et de l'établissement, tout en participant à la transparence administrative ; en gérant le cycle de vie des documents et des données ainsi créés, avec des durées avant élimination ou préservation à long terme précisément définies.

La gestion des documents bureautiques est très représentative de cette évolution. Depuis 2004, elle s'inscrit dans un programme de gestion de l'archivage (*records management*) d'une ampleur considérable, touchant tous les agents de la Bibliothèque et toutes ses activités. Par son biais a été dessinée la trame d'une organisation humaine et technologique permettant la prise en charge de documents nativement numériques dès leur ébauche – courriels compris –, et pouvant accueillir les documents dématérialisés à leur arrivée à la Bibliothèque. Cette chaîne de production documentaire constitue une nouvelle collection au sein du patrimoine numérique de la BnF : ses archives publiques, conservées à terme dans SPAR via la filière « Archives administratives ».

Les extranets

De nombreuses procédures, qu'elles relèvent de la coopération, du dépôt légal ou d'autres activités, impliquent des acteurs externes à la BnF : éditeurs ou déposants pour le dépôt légal, bibliothèques et institutions partenaires, ou encore prestataires, dans le cadre de marchés publics. La BnF a ainsi développé une expertise dans la création de plateformes d'accès réservé dédiées aux échanges avec ces tiers : les extranets.

Le site Extranet du dépôt légal offre un service de déclaration en ligne destiné à faciliter les démarches de dépôt légal des éditeurs et producteurs de documents. Les données bibliographiques sont saisies par les déposants et, le plus souvent, accompagnées des images de couverture. Ces données sont récupérées dans le système d'information de la BnF et les récépissés électroniques sont transmis instantanément en retour. Si l'éditeur donne son accord, les données sont publiées sous forme d'annonces sur le site Nouveautés Éditeurs.

Les extranets jouent un rôle essentiel dans l'évolution de la BnF vers le numérique, en simplifiant et dématérialisant les procédures.

De son côté, l'Espace Coopération, créé en 2011, est un outil de travail qui accompagne au quotidien les partenaires de la BnF, qu'il s'agisse de leurs démarches administratives (signature de conventions…), du paramétrage de la présentation de leurs collections dans Gallica ou de leur adhésion à l'offre de tiers archivage. L'Espace Coopération propose également un lieu d'échanges sous la forme d'un forum et d'un fil d'actualités.

Enfin, des extranets dédiés à des fonctions particulières ont été mis en place : la plateforme PLATON est ainsi utilisée pour le transfert de fichiers numériques entre éditeurs et associations agréées aux fins d'adaptation des ouvrages pour des publics en situation de handicap, au titre de l'exception prévue dans le code de la propriété intellectuelle.

Ces extranets jouent un rôle essentiel dans l'évolution de la BnF vers le numérique. En simplifiant et dématérialisant les procédures, ils ont ouvert la porte à des transformations profondes des modalités du travail dans la Bibliothèque, comme lorsqu'il s'agit d'intégrer automatiquement les métadonnées des éditeurs dans le catalogue. En confortant l'image de modernité et de confiance de la BnF, ils permettent l'entrée de nouveaux objets dans les collections, tels que les livres numériques ou la numérisation rétrospective des bibliothèques partenaires. Enfin, lorsque la collecte par robot de la presse quotidienne régionale, en substituant des supports numériques aux supports analogiques pour la communication, réduit de 300 000 euros par an le budget du microfilmage de conservation, l'automatisation s'avère aussi synonyme d'économies.

Ces tendances devraient conduire à une convergence des interfaces et des outils pour les partenaires de la BnF. On peut ainsi espérer qu'à terme un éditeur qui fournira un fichier numérique pour le dépôt légal, accompagné de ses métadonnées, disposera d'un système d'aiguillage lui permettant d'insérer son document dans le dispositif PLATON pour en permettre l'édition adaptée, d'autoriser la diffusion du résumé et de la couverture dans Nouveautés Éditeurs et dans le catalogue, et d'être assuré de sa conservation pérenne dans SPAR. Les éditeurs qui le souhaiteraient pourraient même, à terme, autoriser une consultation partielle de leurs ouvrages dans Gallica afin d'en augmenter la visibilité.
Ainsi, dans la continuité de l'expérimentation avec les e-distributeurs, qui visait à donner de la visibilité dans Gallica à des livres numériques récents, les éditeurs pourraient se voir proposer de diffuser leurs livres numériques dans Gallica aux mêmes conditions que les ouvrages numérisés dans le cadre du projet ReLIRE, à savoir un accès libre aux premiers 15 % de l'ouvrage avec un renvoi sur les plateformes des libraires, indépendamment de leur diffusion intégrale dans Gallica intra muros.

Les outils de gestion des collections numériques

Pour chaque nouvelle filière numérique, la Bibliothèque se dote d'une couverture optimale de la chaîne de traitement documentaire, de l'entrée des documents à leur diffusion, en passant par leur conservation et leur signalement.

Certains de ces outils, comme ceux qui sont utilisés pour gérer la collection des archives du web, de la collecte à l'accès en salle, ont été conçus et réalisés de manière collaborative, au sein de groupements internationaux de professionnels des bibliothèques qui sont confrontés aux mêmes défis. D'autres, comme CEDN (la chaîne d'entrée des documents numériques), sont construits spécifiquement pour répondre aux besoins de la BnF, en s'appuyant sur des briques existantes de son système d'information.

Élaborer des outils qui permettront, demain, aux bibliothécaires de garder la pleine maîtrise de leurs collections est à la fois un défi et une nécessité.

Une collection numérique est d'autant plus délicate à définir et à appréhender qu'elle entraîne, pour les acteurs qui la manipulent, une perte du rapport concret au document : celui-ci n'est plus un support tangible que l'on peut évaluer, déplacer et stocker ; les tâches associées à la gestion des collections numériques se limitent à une gestion de flux à travers des statistiques ou des fichiers de « logs ». Certaines de ces activités sont désormais effectuées par des informaticiens ou des ingénieurs de production, et non plus des bibliothécaires ou des magasiniers. Élaborer des outils qui permettront, demain, aux acteurs des métiers de bibliothèque de garder la pleine maîtrise de leurs collections est à la fois un défi et une nécessité. C'est à travers des projets comme SIPIL (Système informatique de pilotage) ou le futur outil de publication des collections numériques que la Bibliothèque réunit les indicateurs nécessaires à la conduite des activités numériques au quotidien et invente jour après jour la boîte à outil du nouveau bibliothécaire numérique.

05.A La chaîne de numérisation : processus et système

La gestion des flux de documents à numériser s'appuie sur un système appelé « chaîne d'entrée » qui combine des applications de gestion de la conservation et du magasinage (initialement développées pour des traitements liés aux collections imprimées : désacidification, reliure, etc.), et de nouvelles applications conçues pour gérer les échanges entre la BnF et ses prestataires. et permettre également des contrôles visuels avant mise en ligne.

LES PRINCIPALES ÉTAPES

Un contrôle automatique des documents est mis en œuvre à l'entrée. Vérifiant le format et la cohérence des métadonnées produites, il permet d'identifier les traitements nécessaires à la mise en ligne des documents dans Gallica. C'est au niveau de cette « chaîne d'entrée » que s'opère l'exemplarisation, c'est-à-dire le signalement, dans un catalogue, de la présence d'une version numérisée d'un document, ainsi que diverses transformations d'images pour permettre le zoom, la présentation de vignettes, etc. L'archivage dans SPAR s'y programme aussi, suivant un calendrier échelonné.

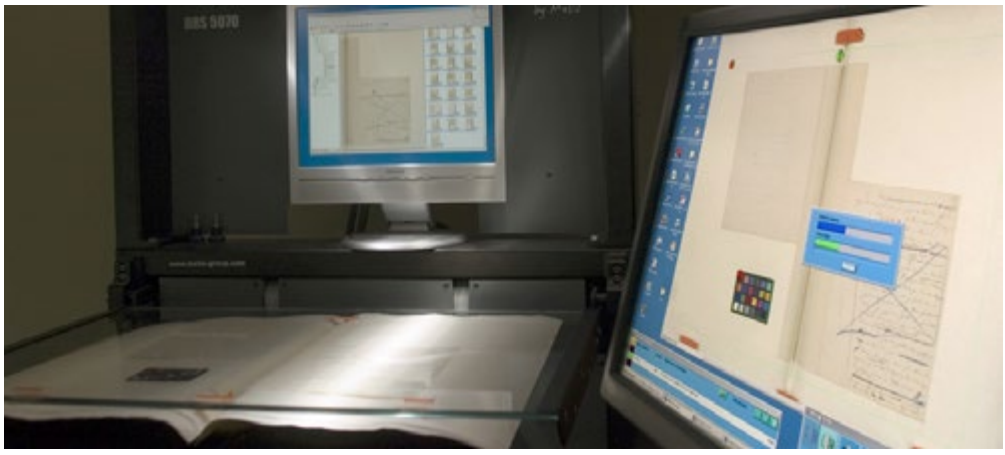
UN SYSTÈME ÉVOLUTIF

Depuis le premier marché passé en 2007, les périmètres et finalités de la numérisation ont considérablement évolué : de nouveaux types de documents ont été ciblés (presse, imprimés précieux issus de la Réserve et de l'Arsenal, manuscrits, estampes, cartes, partitions musicales, objets, etc.), des traitements complémentaires se sont imposés (ajout d'OCR, d'EPUB sur un document déjà numérisé), la filière « urgent lecteur » a été mise en place, et de nombreuses évolutions techniques ont été intégrées (augmentation de la résolution de 300 à 400 DPI, océrisation, documents sonores et vidéos).

LIMITER LES BARRIÈRES TECHNOLOGIQUES

Ces changements vont dans le sens d'un assouplissement de processus devenus trop rigides à l'heure où l'établissement multiplie les partenariats et les projets. Il est également vital de limiter la barrière technologique à l'entrée des marchés de numérisation que peut occasionner une chaîne trop complexe, constituant un frein potentiel pour certains prestataires. La refonte technique de l'application « chaîne de numérisation », initiée en 2015, représente un enjeu stratégique décisif pour la Bibliothèque : il s'agit de saisir cette opportunité pour évaluer fonctionnellement le besoin de souplesse accru qui a émergé depuis neuf ans, et pour adapter les outils de suivi et d'automatisation, dont le rôle est devenu crucial pour gérer la masse des documents numérisés.

La refonte technique de l'application « chaîne de numérisation », initiée en 2015, représente un enjeu stratégique décisif pour la Bibliothèque.



Numérisation d'un manuscrit

05.B
L'extranet du dépôt légal

Le site Extranet du dépôt légal offre un service de déclaration en ligne destiné à faciliter les démarches de dépôt légal des éditeurs et producteurs de documents. Il concerne tous types de documents soumis au dépôt légal auprès de la BnF.

En 2015, l'extranet Editeurs a ainsi été utilisé pour 75 % des dépôts de cartes et plans et 52 % des dépôts de livres imprimés.

UN SERVICE DE DÉCLARATIONS DÉMATÉRIALISÉES

Les données bibliographiques sont saisies par les déposants et, le plus souvent, accompagnées des images de couverture. Ces données sont récupérées dans le système d'information de la BnF et les récépissés électroniques sont transmis instantanément en retour. Si l'éditeur donne son accord, les données sont publiées sous forme d'annonces sur le site Nouveautés Éditeurs. Depuis 2014, s'y ajoute le transfert automatique de métadonnées mis en place avec de grands groupes d'édition : les données descriptives sont transmises chaque semaine par flux automatisés par les distributeurs, qui envoient simultanément les colis de livres imprimés.

LES ATOUTS

Réduisant les délais d'enregistrement des documents et d'expédition des récépissés, l'extranet s'est pleinement adapté aux besoins des éditeurs, fournissant, par ailleurs, des notices enrichies et les identifiants d'auteurs de type ISNI (voir axe 2 fiche F). Il favorise une collaboration étroite entre la BnF et les éditeurs dans le domaine des métadonnées, notamment autour du format ONIX, un format XML utilisé par les professionnels de l'édition à travers le monde et que la Bibliothèque a su à cette occasion s'approprier. En retour, le catalogue de la BnF s'est enrichi de données nouvelles, à forte valeur ajoutée.

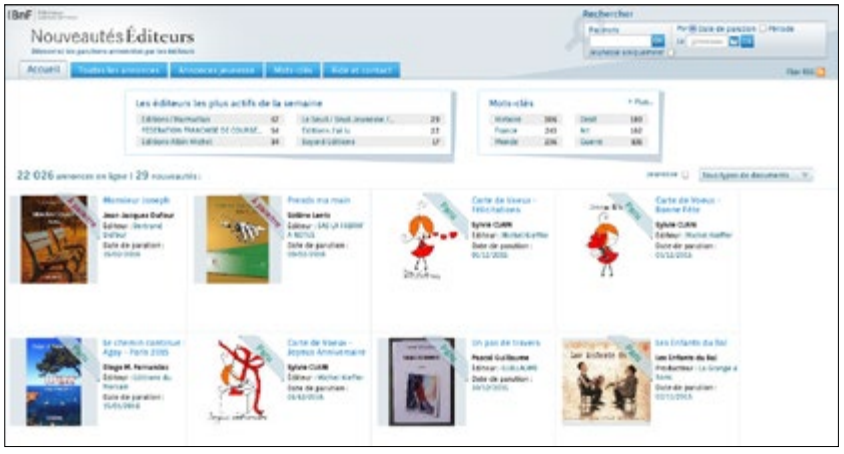
EXTENSION AUX LIVRES NUMÉRIQUES

Les fonctionnalités de l'extranet du dépôt légal vont être étendues pour recevoir le dépôt légal des livres numériques. Les e-distributeurs transmettront par flux les e-books au format EPUB ou PDF, sans mesures de protection (DRM), accompagnés de leurs métadonnées au format ONIX. Après contrôle automatique des fichiers, l'extranet du dépôt légal assurera la récupération des métadonnées, la publication des annonces, le cas échéant, sur Nouveautés Éditeurs, l'enregistrement du dépôt des documents et la création des notices dans le Catalogue général de la BnF (voir axe 1 fiche H).

L'extranet du dépôt légal évoluera ensuite pour permettre le dépôt légal des documents sonores dématérialisés par flux : l'automatisation de la récupération des fichiers numériques et de leurs métadonnées devrait permettre à la Bibliothèque de faire entrer dans ses collections plus de 170 000 albums de musique édités uniquement en ligne, qui échappent actuellement au dépôt légal audiovisuel. Le dépôt des fichiers numériques à l'unité sera également envisagé : les déposants pourront joindre eux-mêmes leurs fichiers aux déclarations saisies en ligne.

En 2015, l'extranet Editeurs a été utilisé pour 75 % des dépôts de cartes et plans et 52 % des dépôts de livres imprimés.

Le site Nouveautés Éditeurs, alimenté par l'extranet du dépôt légal



05.C L'extranet Espace Coopération

Le réseau de partenaires de la BnF (bibliothèques, éditeurs, distributeurs en ligne) s'est densifié au fil des ans, sur tout le territoire français ; il s'est aussi complexifié, avec la mise en place de partenariats à objets de coopération multiples. Un outil de communication, d'échange de documents et de gestion des collections des partenaires dans Gallica était indispensable pour la BnF, mais aussi pour l'ensemble des partenaires afin de simplifier et d'accélérer la mise en commun des ressources documentaires des bibliothèques françaises.

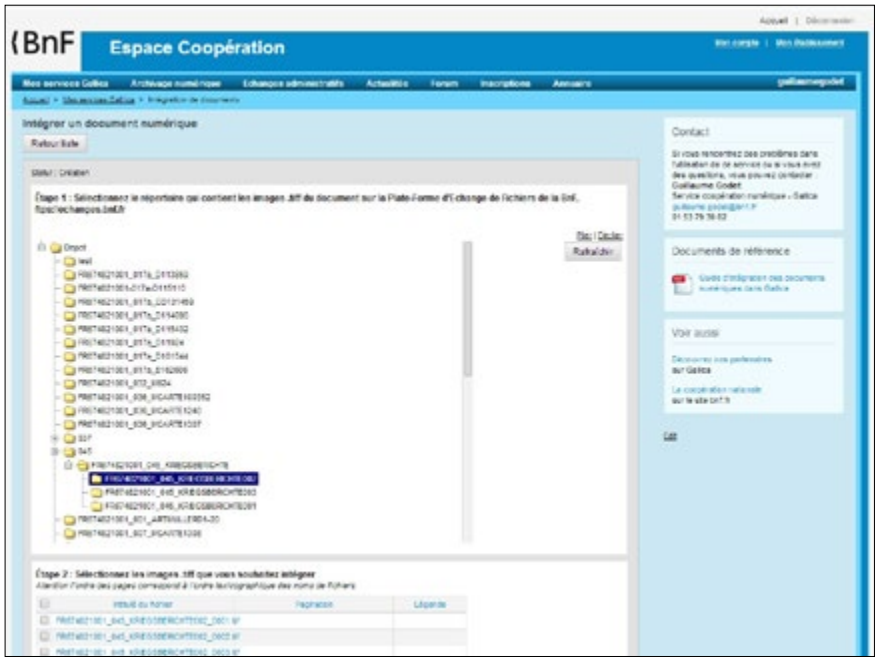
UN OUTIL DE TRAVAIL

Institué en 2011, l'Espace Coopération est un outil de travail qui accompagne au quotidien les partenaires de la BnF ainsi que les clients du tiers archivage, simplifiant leurs démarches et leur permettant de s'informer, voire d'échanger sur le forum. La rubrique Mes services Gallica permet notamment le paramétrage des entrepôts OAI-PMH des partenaires pour le signalement des documents dans Gallica, le suivi des documents transmis à la BnF dans le cadre du marché de numérisation des imprimés, ou encore la création et le suivi des fiches décrivant l'état physique des documents… La rubrique tiers archivage, quant à elle, inclut le suivi des collections conservées par la BnF dans le cadre de l'offre « BnF Archivage numérique » ainsi que la recherche de documents et le suivi des demandes de communication.

L'un des développements majeurs de l'Espace Coopération vise à en faire le principal point d'entrée pour l'intégration de documents numérisés par des partenaires.

LES AXES DE DÉVELOPPEMENT

L'un des développements majeurs de l'Espace Coopération, en cours et à venir, vise à en faire le principal point d'entrée pour l'intégration de documents numérisés par des partenaires, soit à des fins de diffusion dans Gallica (voir axe 4 fiche J), soit en vue de la création d'un site Gallica marque blanche (voir axe 4 fiche F). Pour les années suivantes, l'Espace Coopération devra développer de nouveaux services en accord avec la politique de coopération de la BnF et le projet Gallica marque blanche, pour permettre aux partenaires de gérer au mieux la sélection, la production, le suivi, la diffusion et l'archivage de leurs collections numériques, ainsi que les échanges administratifs avec la BnF.



Intégration de fichiers numériques des partenaires en vue de leur consultation dans Gallica et dans les sites en marque blanche

05.D
La plateforme PLATON pour l'exception handicap

La loi de 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information a institué une exception au bénéfice des personnes atteintes d'un handicap. La reproduction des œuvres sur des supports adaptés aux publics handicapés peut être effectuée librement et sans contrepartie financière par des associations ou des établissements habilités conjointement par le ministre chargé de la culture et le ministre chargé des personnes handicapées.

UNE PLATEFORME DÉDIÉE

Afin d'exercer son rôle de tiers de confiance dans l'application de l'exception handicap prévue dans le code de la propriété intellectuelle, la BnF a développé la plateforme PLATON (Plateforme de transfert des ouvrages numériques) qui permet aux organismes agréés d'effectuer des demandes et aux éditeurs de déposer leurs fichiers. Les organismes accèdent ensuite aux fichiers et peuvent déclarer les adaptations, en lien avec la Banque de données de l'édition adaptée (BDEA), administrée par l'Institut national des jeunes aveugles, grâce à une solution d'interopérabilité développée à cet effet.

UN IMPACT TRÈS POSITIF

Depuis son lancement en 2010, PLATON a permis de développer significativement l'offre accessible aux personnes handicapées, principalement les déficients visuels : plus d'une quarantaine d'organismes agréés et plus de 700 éditeurs utilisent ce service, réalisant en moyenne 6 000 à 8 000 adaptations par an. L'abonnement en flux numérique à une importante sélection de revues et de périodiques, via deux agrégateurs de presse, est par ailleurs possible sur PLATON depuis 2013. La « Rentrée littéraire accessible » est organisée depuis septembre 2013, l'édition 2015 ayant mis à disposition du public empêché 330 nouveautés littéraires grâce au soutien du CNL et à une collaboration exemplaire des acteurs du dispositif, soit 42 % de plus qu'en 2014.

Depuis son lancement en 2010, PLATON a permis de développer significativement l'offre de livres et revues accessibles aux personnes handicapées.

D'autres évolutions devraient encore amplifier l'impact de PLATON à l'avenir, notamment le dépôt des fichiers adaptés, assorti de la communication facilitée desdits fichiers aux organismes agréés. Dans un avenir proche, les fichiers des manuels scolaires devraient faire l'objet d'un dépôt systématique, permettant aux élèves en situation de handicap d'en disposer dans les meilleurs délais. Enfin, la BnF envisage de proposer aux éditeurs, sur la base du volontariat, de corréler à terme le dépôt légal des livres numériques à la mise à disposition des fichiers dans PLATON.

Un processus sécurisé

Les organismes agréés ont la possibilité de demander, dans les dix ans suivant le dépôt légal des œuvres imprimées et à condition que ce dépôt légal soit postérieur à 2006, que les fichiers numériques ayant servi à l'édition de ces œuvres soient mis à leur disposition par les éditeurs via la plateforme PLATON.

La BnF agit dans ce dispositif comme un tiers de confiance. Une fois l'adaptation effectuée (adaptation en gros caractères, braille, édition numérique adaptée…), les fichiers sources sont détruits par les organismes d'adaptation, mais conservés sans limitation de date à la BnF.

05.E La gestion électronique des documents

La BnF a commencé à utiliser des documents nativement numériques dès le développement de l'informatique bureautique dans les administrations françaises. C'est toutefois en 2004 que la nécessité de passer d'une gestion individuelle des documents à une organisation plus collective s'est imposée, face à la multiplication de projets transverses associant de nombreuses compétences internes.

LES PREMIERS PAS

Le programme de gestion de l'archivage (*records management)* a dessiné la trame d'une organisation humaine et technologique permettant la prise en charge de documents nativement numériques dès leur ébauche – courriels compris –, et pouvant accueillir les documents dématérialisés à leur arrivée à la Bibliothèque. Ainsi a été instituée en 2006 une chaîne de production documentaire allant de la création de données jusqu'à la fin du cycle de vie de l'information, l'ensemble constituant les archives publiques produites par la BnF, en partie conservées dans SPAR via la filière « Archives administratives ».

COMBINER INTÉRÊTS STRATÉGIQUES ET OPÉRATIONNELS

Une production documentaire efficace consiste à combiner les intérêts stratégiques de l'établissement et le bon fonctionnement quotidien des activités individuelles. À travers elle se préservent et se transmettent le capital de connaissances de l'institution et sa mémoire. Savoirs, expériences, informations y sont consignés : bien décrits et correctement classés, les documents créés sous forme électronique sont aisément exploitables (selon les droits d'accès accordés). Leur validité administrative et juridique est par ailleurs mieux établie, au bénéfice des personnes et de la transparence administrative.

ACCOMPAGNER LE CHANGEMENT

En cours de déploiement, un dispositif actualisé de gestion électronique des documents améliore l'ergonomie et l'intuitivité des interfaces, ainsi que les possibilités d'autoformation de l'utilisateur. Les caractéristiques fondamentales du système restent pour leur part inchangées : un plan de classement, une liste de typologies associées à une bibliothèque de modèles, un catalogue documentaire fédéré, un identifiant pérenne et unique pour chaque document et dossier, attribué automatiquement. Par souci de simplicité, toutes les métadonnées pouvant être récupérées de façon automatisée le sont, permettant aux rédacteurs de se concentrer sur l'essentiel : leur activité courante.

Savoirs, expériences, informations sont consignés : bien décrits et correctement classés, les documents créés sous forme électronique sont aisément exploitables.

La mise en place d'une telle organisation a exigé un important effort de formation et d'adaptation de la part des 2 400 agents de la Bibliothèque. Il leur a ainsi fallu apprendre dans quel espace ou dossier enregistrer tel ou tel document, quelles règles de nommage utiliser, comment transmettre l'information en interne ou hors de la Bibliothèque. Le bilan de cet accompagnement au changement s'avère très positif, avec une adhésion de la grande majorité des agents et l'intégration des équipes RM (*records management*) au sein des projets stratégiques de la BnF. Il convient à présent d'améliorer encore l'ergonomie des outils, de poursuivre l'intégration dans les processus métiers et de développer le conseil et l'accompagnement des équipes.

05.F
Les outils de l'archivage du web

La mise en œuvre, entre 2006 et 2013, d'un circuit complet d'archivage du web a supposé de développer ou d'intégrer un panel d'outils spécifiques couvrant les différentes étapes de la chaîne. Constituée suivant une approche modulaire et dans le respect des standards internationaux, cette chaîne repose dans la mesure du possible sur des logiciels libres, le plus souvent développés au sein du Consortium international pour la préservation de l'internet (IIPC).

SÉLECTIONNER

BnF Collecte du web ou BCWeb, l'outil développé par la BnF pour sélectionner les sites à collecter, a évolué vers un projet de logiciel libre impliquant de multiples acteurs. Les informations qu'il contient sont réutilisées à des fins de signalement, notamment par l'intermédiaire de data.bnf.fr.

COLLECTER

La BnF utilise le robot d'archivage Heritrix, lui-même piloté par un outil de gestion de flux, NetarchiveSuite, qui permet de programmer les collectes, d'en surveiller le déroulement et d'en assurer le contrôle qualité. La mise à jour technologique de ces outils, en vue d'améliorer notamment la prise en compte des plateformes de diffusion audiovisuelles, des réseaux sociaux ou encore des applications, est communément portée par les membres d'IIPC.

SIGNALER ET DONNER ACCÈS

La BnF exploite également les outils mutualisables (le logiciel SOLR pour l'indexation en texte intégral, le format WAT – développé par Internet Archive – pour l'analyse), et s'implique par ailleurs dans le projet Open Wayback utilisé pour donner accès aux archives en salles de lecture.

PRÉSERVER

Le système SPAR (voir axe 2 fiche B) a vocation à recevoir la totalité des collections du dépôt légal de l'internet. En parallèle, une réflexion sur les stratégies de préservation est menée par le groupe de travail préservation d'IIPC, qui s'est lancé dans une opération de documentation du web, afin de pouvoir reconstituer, à l'avenir, l'environnement de travail d'un internaute en 2000, 2005 ou 2010.

L'outil BnF Collecte du web, pour la sélection des sites

Accès aux archives de l'internet en salle de lecture

05.G La qualité

Opérer un suivi de la qualité des collections numériques est vital si l'on veut s'assurer de l'adéquation entre la nature des ressources produites ou collectées et les attentes nourries par les professionnels comme par le public à l'égard de ces ressources. Exprimée sous forme d'indicateurs statistiques, la qualité devient l'un des outils de pilotage de la stratégie numérique de la Bibliothèque, suivi avec d'autant plus d'attention qu'il représente l'un des marqueurs identitaires d'une bibliothèque nationale.

DES CRITÈRES ADAPTÉS SUIVANT LE CONTEXTE

La définition et l'atteinte d'objectifs de qualité ne vont pas de soi, car la qualité n'est pas une valeur absolue. À titre d'exemple, dans une perspective de préservation, la qualité d'un document numérisé est appréciée à l'aune de sa fidélité à l'égard du document original. Dans une perspective de consultation, au contraire, un document fortement retouché peut paraître de meilleure qualité qu'un document respectant strictement les défauts de l'original.

Par ailleurs, la qualité est parfois gouvernée par des impératifs d'ordre technique ou contractuel. Ainsi, dans le cas d'une collecte web, la couverture d'un site n'est pas la même suivant que l'on effectue une collecte de surface (page d'accueil du site uniquement) ou une collecte en profondeur (toutes les pages et tous les médias du site). Évaluer la qualité de la collecte ne consistera pas, dès lors, à vérifier que le site web a été entièrement couvert, mais bien qu'il a été couvert dans les formes et périmètres souhaités.

LE DÉFI DES GRANDS VOLUMES

S'agissant des marchés de numérisation, la volumétrie traitée et la mise en place d'un flux tendu entre la BnF et le prestataire ont rendu impossible le contrôle à l'unité de la qualité des documents. D'autres méthodes ont dû être élaborées : phases de test et de galop d'essai plus soutenues, audit des chaînes du prestataire, contrôle par carottage.

L'outil de contrôle qualité de la chaîne de numérisation (incunable de la Réserve des livres rares, Vélins 155)

L'ÉPREUVE DU TEMPS

Il peut arriver qu'un défaut soit détecté des années après une numérisation. Par ailleurs, les attentes et les méthodes évoluent : une numérisation ancienne, jugée en son temps de bonne qualité, pourra apparaître insatisfaisante plus tard, justifiant une nouvelle opération de numérisation (remplacement du noir et blanc par des niveaux de gris ou par la couleur, océrisation plus avancée).

L'ENJEU DES MODES DE PRODUCTION COLLABORATIFS

De surcroît, le niveau de qualité fixé par la BnF pour ses propres collections numérisées peut être, en l'état, un obstacle au développement de modes de production alternatifs, voire collaboratifs. Partager l'effort de production nécessite d'accepter que des niveaux de qualité différents coexistent dans les collections nationales, signalés comme tels.

Ce qui est vrai pour la collection l'est également des métadonnées. Envisager une production collaborative des notices (avec les éditeurs, avec le réseau OCLC, avec d'autres acteurs comme l'ABES…) suppose d'adapter l'exigence de qualité ou d'accepter la coexistence de niveaux de qualité différenciés entre les données.

Enfin, la notion de qualité ne doit pas s'arrêter à la seule gestion des collections : mesurer la qualité du service offert à l'utilisateur final, en termes de contenus, de délai de réponse ou encore de continuité de service s'avère tout aussi crucial.

Partager l'effort de production nécessite d'accepter que des niveaux de qualité différents coexistent dans les collections nationales.

05.H
Les outils de gestion des collections numériques

L'outillage des tâches de gestion des collections numériques dans toute leur diversité ne peut se résumer à une seule application qui couvrirait l'ensemble des besoins. Il s'agit au contraire de réunir une panoplie d'outils, chacun traitant de manière spécifique une activité donnée : suivi des marchés, conservation, accès en consultation, etc. Certaines filières comme l'archivage de l'internet ou le système audiovisuel (SA, voir axe 2 fiche E) disposent de leur propre panel d'outils pour sélectionner, collecter, traiter et diffuser les collections correspondantes.

BÉNÉFICIER D'INDICATEURS

S'agissant de la numérisation, le suivi de la production s'opère à l'aide de l'outil de requêtage SIPIL (Système informatique de pilotage) conçu pour collecter et interroger un grand nombre de données au sein d'ensembles larges, ou bien document par document. Anomalies et rejets peuvent ainsi être identifiés, puis traités. Par ailleurs, à partir des données qu'il fouille, l'outil SIPIL permet de produire de précieux indicateurs (hebdomadaires, mensuels). Ces indicateurs servent en premier lieu à apprécier la performance et la stabilité des opérations de numérisation ; mais les besoins de suivi ont progressivement évolué, se décentrant des marchés de numérisation pour viser une meilleure connaissance des collections numériques : évaluer les besoins de traitements, mieux sélectionner les fonds à numériser ou à retraiter, supprimer les doublons, identifier ceux présents dans Gallica ou absents de SPAR, notamment.

AGIR SUR LES DONNÉES

Ce travail de description des collections numérisées pousse SIPIL à satisfaire des besoins de plus en plus larges. Il s'agit de rassembler et mettre en cohérence des données issues de différents silos : les données d'entrées, dont éventuellement celles liées aux demandes de numérisation, les données des magasins numériques et les données du catalogue.

Cependant, SIPIL est un outil statique ; il ne permet pas d'agir sur les données ou sur les collections numériques pour les traiter. Ce type d'opérations requiert d'autres dispositifs, dont l'outil de publication, qui doit permettre d'aiguiller les documents numériques vers les différentes interfaces de consultation (Gallica, Gallica intra muros, marques blanches…), est un bon exemple. Les outils de gestion des collections numériques doivent permettre aussi bien un pilotage « macro », appliqué aux grandes masses, qu'une gestion opérationnelle de niveau « micro », tel le suivi précis d'un document dans le processus de numérisation, d'entrée, d'archivage et d'accès. L'un des enjeux est de faire le lien entre les collections physiques et les collections numériques, afin de les considérer comme un tout aussi bien du point de vue de la conservation que de celui de l'accès. Il importe de faire en sorte que ces outils deviennent plus concrets, intuitifs et conviviaux pour les agents qui les utilisent quotidiennement. La création et l'amélioration continue des outils doivent rester une priorité pour la Bibliothèque.

Ces outils doivent permettre aussi bien un pilotage « macro », appliqué aux grandes masses, qu'une gestion opérationnelle de niveau « micro ».

L'outil de publication des documents numériques

Cet outil, dont le développement a débuté en 2015, a vocation à permettre un pilotage plus fin de la diffusion des contenus de la bibliothèque numérique.

Il s'agit d'aiguiller chaque document vers l'environnement de consultation approprié : un document libre de droits destiné à la consultation sur internet apparaîtra dans Gallica, tandis qu'un livre numérique entré par dépôt légal verra sa consultation limitée à la bibliothèque de recherche.
Dans certains cas, l'accès à la collection doit aussi être géré suivant des critères bibliothéconomiques : on pourra retirer de la consultation un document numérisé lorsqu'une version de meilleure qualité sera disponible, par exemple, fonction qui s'apparente au désherbage.

05.1 Les acquisitions et dons de documents numériques (ADDN)

L'arrivée à la BnF de documents sous forme numérique par d'autres voies que celle du dépôt légal constitue une nouvelle occasion pour l'établissement de repenser ses pratiques en matière de conservation, de catalogage et de communication. C'est un défi, car le nombre de documents en attente s'élève à plusieurs dizaines de milliers.

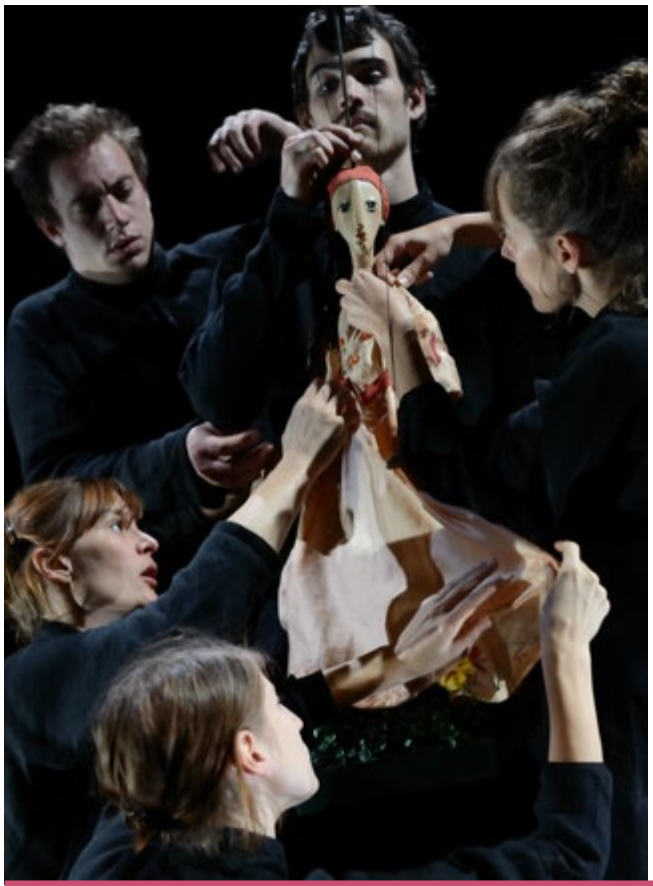
DES DOCUMENTS DE NATURE PATRIMONIALE

Introduits par don ou par acquisition, ces documents numériques sont de nature patrimoniale (archives, niches éditoriales), et ont vocation à être conservés de façon pérenne. Relevant du domaine français sans être couverts par le dépôt légal, ils n'existent cependant pas sur support physique et illustrent la bascule complète de certains secteurs vers le numérique.

Le don, en 2004, par lequel l'écrivain Pierre Guyotat remit au département des manuscrits ses archives et correspondances électroniques (courriels, textos) a été l'un des déclencheurs d'une réflexion sur la nécessité de s'adapter aux nouvelles pratiques des déposants. Aussi la prise en compte des documents numériques dans les filières d'entrée constitue-t-elle un enjeu capital pour la Bibliothèque.

UNE NOUVELLE FILIÈRE

La variété des formats existants et la diversité des types de documents collectés (livres, articles, périodiques, photographies, cartes, partitions, archives privées…) constituent les principales difficultés du projet. S'y ajoutent la gestion des droits d'accès aux documents et la variété des situations en la matière, obligeant à une gestion très fine de chaque contenu. À cette fin, une nouvelle filière, Acquisitions et dons de documents numériques (ADDN), couvrant toutes les étapes du traitement, a été mise en place. La réalisation du projet s'appuie sur des briques logicielles existantes ou en cours de création dans le cadre d'initiatives menées concomitamment : dépôt légal du livre numérique, Gallica intra muros, ou encore SPAR.



Brigitte Pougeoise, « Le conte d'hiver », Compagnie Arketal, Chartreuse du Val-de-Bénédiction, Villeneuve-lès-Avignon, 2012

Ces documents n'existent pas sur support physique et illustrent la bascule complète de certains secteurs vers le numérique.

Un exemple de collection pilote : les arts du spectacle

ADDN permettra dans un premier temps de traiter des photographies numériques natives, acquises auprès de photographes de spectacle. Parmi ces ensembles figurent le corpus Joël Verhoustraeten sur le spectacle de rue (2 750 fichiers au format JPEG) ainsi que le fonds Brigitte Pougeoise, photographe de spectacle de marionnettes (1 070 fichiers aux formats JPEG et TIFF). Ces documents numériques natifs s'inscrivent dans la continuité de la politique d'acquisition de photographies, visant à conserver la mémoire du spectacle vivant.

BnF

06

ENGAGER UNE DÉMARCHE D'INNOVATION DURABLE

72
formations au numérique en 2013

611
personnes formées en 2013

16 millions €
consacrés au développement logiciel
dans le domaine numérique
entre 2011 et 2014

Jusqu'au milieu des années 2000, le numérique représentait pour la BnF un terrain d'exploration et d'expérimentations diverses, qui a vu la montée en puissance de Gallica et la naissance de SPAR ainsi que des archives du web.

L'établissement a ensuite connu une phase d'implantation du numérique dans l'organisation, qui a progressivement touché tous les services et tous les métiers : la Bibliothèque, à l'instar de tous les protagonistes de la sphère culturelle, a connu sa transition numérique.

Cette mutation profonde, qui n'est pas encore achevée, n'aurait pas été possible si la Bibliothèque n'avait mobilisé ses moyens et ses compétences au service d'une dynamique de changement qu'elle a su maintenir sur plusieurs années.

ZOOM SUR QUELQUES POINTS FORTS

06.A_ De nouveaux formats pour le livre numérique : EPUB, DAISY
06.B_ Recherche et développement autour de l'océrisation

MOYENS, OUTILS ET MÉTHODES

06.C_ Les moyens humains
06.D_ L'investissement dans le développement informatique
06.E_ Communication interne et accompagnement au changement
06.F_ La formation continue

AXES DE DÉVELOPPEMENT

06.G_ La fouille de textes et de données (*text & data mining*)
06.H_ La numérisation en 3D
06.I_ Le projet de coopération avec le CEA

Expérimenter et innover

La numérisation produit un matériau fondamental pour la recherche scientifique. En encourageant le développement d'outils innovants et toujours plus performants, en matière de reconnaissance optique de caractères, par exemple, elle favorise les recherches, notamment dans le domaine des sciences humaines et sociales. Mis en lumière à l'occasion de leur numérisation, des corpus complets deviennent autant de terrains d'étude. De plus, la masse des données numériques disponibles atteint aujourd'hui un tel volume que seul un traitement informatisé permet de l'appréhender dans son ensemble. Le rapport aux documents s'en trouve révolutionné : fouille de textes, de données et d'URL font émerger de nouveaux corpus et de nouvelles méthodes de recherche. En investissant ce champ d'étude, la BnF entend prendre part au développement des humanités numériques en France.

Fouille de textes, de données et d'URL font émerger de nouveaux corpus et de nouvelles méthodes de recherche. En investissant ce champ d'étude, la BnF entend prendre part au développement des humanités numériques en France.

La BnF participe à différents projets conçus autour des systèmes innovants de production de l'information numérique. Plusieurs orientations stratégiques sont privilégiées, de l'analyse des médias visuels avec le laboratoire ETIS (université de Cergy) à la reconnaissance d'entités nommées (personnes, lieux, dates, concepts…) faisant le lien entre les entités présentes dans un document et les référentiels BnF ou autres. Les procédés de signalement et de recherche évoluent comme dans le cas de CartoMundi, à la fois catalogue collectif et outil de visualisation de 2 000 séries cartographiques françaises et étrangères. L'amélioration de l'OCR passe par la production participative de contenus, ou *crowdsourcing* (projet de recherche OZALID piloté par Orange, qui a débouché sur la création de la plateforme expérimentale Correct). On peut également citer la modélisation et la réutilisation des données dans le web sémantique ou encore l'analyse des corpus archivés dans le cadre du dépôt légal de l'internet.

Les projets de recherche et de développement (R&D) doivent être distingués des projets d'innovation : alors que l'innovation débouche sur un produit opérationnel (data.bnf.fr, par exemple), la R&D a pour objectif d'expérimenter des dispositifs à usage unique ou temporaire (prototypes). Il s'agit donc d'un véritable travail de recherche, le plus souvent piloté par des chercheurs professionnels (laboratoires, universités…) auxquels la BnF associe experts et équipes opérationnelles. Elle perçoit d'ailleurs des financements à cette fin, dans le cadre d'appels à projets (Agence nationale de la recherche, projets européens…). Ces prototypes, s'ils n'ont pas vocation à être réimplantés tels quels dans les applications du système d'information de la Bibliothèque, génèrent un réel gain de compétences et aident à valider les hypothèses de développement stratégique de l'établissement. La R&D permet ainsi d'expérimenter à échelle réduite des dispositifs, d'évaluer leur succès et leur pertinence en amont d'investissements plus durables.

Car la constitution de corpus numériques et la numérisation du patrimoine à des fins de recherche ne sauraient être conduites indépendamment d'une interrogation sur les usages. Face à l'éclatement des pratiques numériques, capter et analyser ces usages – et notamment les modalités d'appropriation des collections numériques de Gallica par les Gallicanautes – constituent aujourd'hui un véritable défi. Le partenariat avec Télécom ParisTech autour du Laboratoire d'étude des usages du patrimoine en ligne des bibliothèques (Bibli-Lab) a été créé à cette fin (voir axe 3 fiche G). Sans préjuger des enseignements que livreront ces travaux, on peut penser qu'ils aideront sans doute à imaginer, dans un avenir proche, toutes les informations que les collections numériques sont en mesure de livrer à l'aide des outils de fouille de données et d'analyse des mégadonnées (*big data*).

Les moyens du changement

Les compétences acquises et les expérimentations menées dans le cadre de la recherche se traduisent, dans les activités courantes de la Bibliothèque, par une évolution de la culture de l'établissement. En se tournant vers le numérique, la BnF a pris acte de la nécessité d'adapter constamment ses services aux usages émergents ainsi qu'aux possibilités nouvelles offertes par la technologie.

La Bibliothèque s'est ainsi dotée d'outils de gouvernance destinés à fixer et maintenir le cap du changement, en évitant l'effet de déstabilisation parfois induit par le mouvement perpétuel qui caractérise la transition numérique. Ont été créés des comités (numérisation, conservation, pilotage des projets européens…), des référentiels (charte de la politique documentaire, charte de la conservation, référentiels techniques à destination des agents, des partenaires et des prestataires…), des outils de pilotage et de suivi (SIPIL – Système informatique de pilotage de la bibliothèque, voir axe 5 fiche H), des dispositifs d'évaluation et d'observation des pratiques et des usages (observatoire ORH-ION sur l'évolution des pratiques métier en interne, Bibli-Lab sur les usages des internautes, voir axe 3 fiche G).

Être toujours en mouvement requiert des investissements considérables, notamment sur le plan informatique. Exigeant de bien fédérer les compétences des bibliothécaires et celles des informaticiens, ces mutations profondes ont conduit l'établissement à redéfinir ses méthodologies de travail. Le passage au développement en méthode agile a ainsi été un facteur clé de l'innovation numérique, se traduisant par une conception résolument centrée sur les besoins des utilisateurs, un partage d'expertise réciproque entre connaissance des collections et maîtrise des technologies, et une souplesse accrue dans la planification des développements.

Le facteur humain joue également un rôle décisif dans la transition numérique. La gestion prévisionnelle des effectifs, des emplois et des compétences (GPEEC), démarche par laquelle la BnF analyse l'évolution dans le temps de ses ressources humaines en prenant en compte les compétences et l'évolution des fonctions, intègre pleinement le numérique. En la matière, les initiatives destinées à accompagner la transition sont variées : attribution de ressources destinées à faire émerger de nouveaux projets, opérations de sensibilisation en nombre d'agents pour qui le numérique n'est pas au cœur du quotidien… L'innovation gagne aussi les méthodes choisies pour conduire cet accompagnement : la formation se réinvente dans des ateliers où les experts partagent leur savoir-faire avec leurs collègues, la communication interne envisage sa transition vers le collaboratif à travers l'idée d'un réseau social d'entreprise.

La BnF a entrepris de se rapprocher de laboratoires d'innovation afin de mettre en présence des exemples de transition numérique pris hors de la sphère des bibliothèques et les gisements d'expertise et d'imagination de la Bibliothèque.

Afin de stimuler davantage encore cette dynamique, la BnF est membre depuis 2015 d'Ideas Laboratory, une entité d'innovation ouverte qui regroupe plusieurs structures sous le pilotage du Commissariat à l'énergie atomique, où chacun questionne ses pratiques selon une logique à la fois pragmatique et exploratoire. L'installation à la BnF d'une annexe parisienne de ce laboratoire est l'occasion de mettre en présence des exemples de transition numérique pris hors de la sphère des bibliothèques et la richesse que constituent les gisements d'expertise et d'imagination, souvent cachés, de toutes les strates de la Bibliothèque. Ce qui se joue ici est un véritable changement de visage pour l'établissement qui devra, au tournant du quart du siècle, réinventer les modalités de mise en œuvre de ses missions dans un monde où l'accès à la culture et au savoir connaît une véritable révolution.

06.A De nouveaux formats pour le livre numérique : EPUB, DAISY

L'EPUB s'est imposé comme le standard de référence pour la production et la diffusion de livres numériques. Ce format universel répond aux exigences de mobilité, de portabilité et d'accessibilité, conformément à l'évolution des pratiques culturelles contemporaines. Dans sa dernière version (EPUB 3), ce format satisfait tous les usages de lecture (publics voyants, malvoyants, non-voyants, dyslexiques, etc.), pour tout type de contenu (documentaire, littérature, jeunesse, scolaire, universitaire).

LE RESPECT DE L'ÉCOSYSTÈME DU LIVRE NUMÉRIQUE

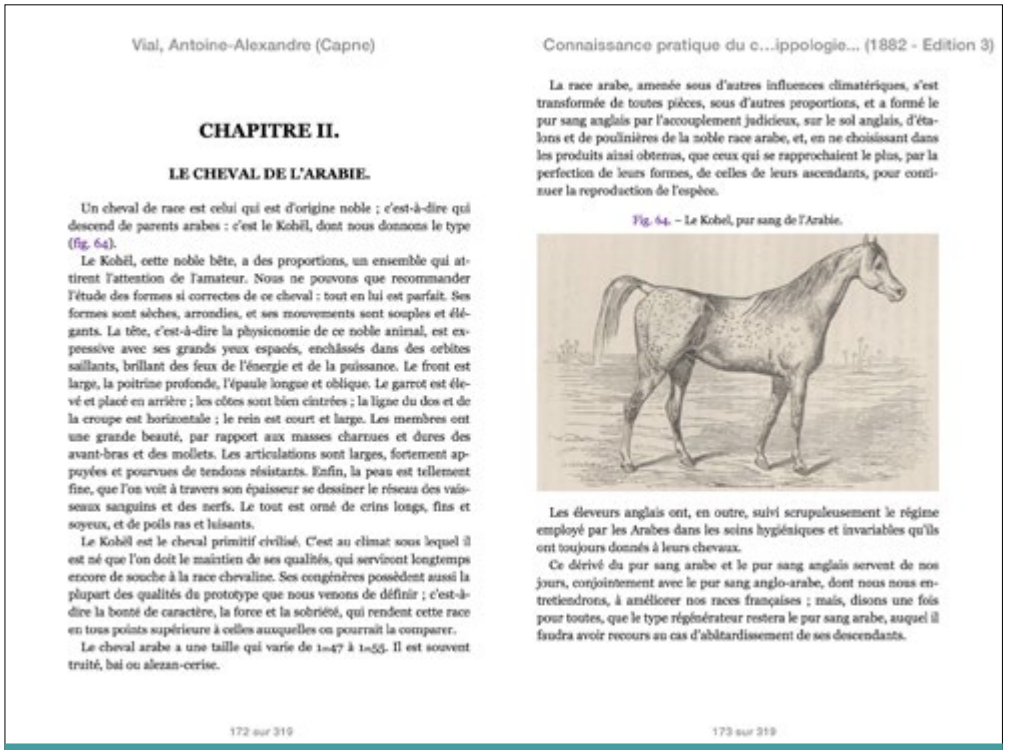
Depuis 2013, la production de livres numériques en EPUB fait partie intégrante de la stratégie de diffusion de Gallica. Cependant, à l'heure où un vaste écosystème numérique se met en place, la BnF veille à ce que sa production d' EPUB gratuits pour Gallica ne soit pas perçue comme concurrente de l'offre commerciale. C'est pourquoi le développement de cette offre se fait à un rythme constant mais mesuré : 3 200 EPUB seront produits sur la période 2014-2017. Sont privilégiés les auteurs rares, les critiques et commentaires des œuvres du domaine public, les entreprises savantes, ou encore des corpus thématiques.

UN FORMAT ACCESSIBLE

Le format EPUB n'est pas la seule technologie adoptée par la BnF pour diffuser plus largement ses collections numérisées. En produisant également des fichiers au format DAISY, destiné à l'accessibilité aux personnes en situation de handicap, l'établissement favorise l'accès de tous aux œuvres du patrimoine écrit, tout en développant encore davantage sa maîtrise technique des nouveaux formats.

La production de livres numériques en EPUB fait partie intégrante de la stratégie de diffusion de Gallica.

Un ouvrage de Gallica au format EPUB



Vial, Antoine-Alexandre (Capne)

CHAPITRE II.

LE CHEVAL DE L'ARABIE.

Un cheval de race est celui qui est d'origine noble ; c'est-à-dire qui descend de parents arabes : c'est le Kohël, dont nous donnons le type (fig. 64).

Le Kohël, cette noble bête, a des proportions, un ensemble qui attirent l'attention de l'amateur. Nous ne pouvons que recommander l'étude des formes si correctes de ce cheval : tout en lui est parfait. Ses formes sont sèches, arrondies, et ses mouvements sont souples et élégants. La tête, c'est-à-dire la physionomie de ce noble animal, est expressive avec ses grands yeux espacés, enchâssés dans des orbites saillants, brillant des feux de l'énergie et de la puissance. Le front est large, la poitrine profonde, l'épaule longue et oblique. Le garrot est élevé et placé en arrière ; les côtes sont bien cintrées ; la ligne du dos et de la croupe est horizontale ; le rein est court et large. Les membres ont une grande beauté, par rapport aux masses charnues et dures des avant-bras et des mollets. Les articulations sont larges, fortement appuyées et pourvues de tendons résistants. Enfin, la peau est tellement fine, que l'on voit à travers son épaisseur se dessiner le réseau des vaisseaux sanguins et des nerfs. Le tout est orné de crins longs, fins et soyeux, et de poils ras et luisants.

Le Kohël est le cheval primitif *civilisé*. C'est au climat sous lequel il est né que l'on doit le maintien de ses qualités, qui servirent longtemps encore de souche à la race chevaline. Ses congénères possèdent aussi la plupart des qualités du prototype que nous venons de définir ; c'est-à-dire la bonté de caractère, la force et la sobriété, qui rendent cette race en tous points supérieure à celles auxquelles on pourrait la comparer.

Le cheval arabe a une taille qui varie de 1m47 à 1m55. Il est souvent truité, bai ou alezan-cerise.

172 sur 319

Connaissance pratique du c...ippologie... (1882 - Edition 3)

La race arabe, amenée sous d'autres influences climatériques, s'est transformée de toutes pièces, sous d'autres proportions, et a formé le pur sang anglais par l'accouplement judicieux, sur le sol anglais, d'étalons et de poulinières de la noble race arabe, et, en ne choisissant dans les produits ainsi obtenus, que ceux qui se rapprochaient le plus, par la perfection de leurs formes, de celles de leurs ascendants, pour continuer la reproduction de l'espèce.

Fig. 64. – Le Kohel, pur sang de l'Arabie.

Les éleveurs anglais ont, en outre, suivi scrupuleusement le régime employé par les Arabes dans les soins hygiéniques et invariables qu'ils ont toujours donnés à leurs chevaux.

Ce dérivé du pur sang arabe et le pur sang anglais servent de nos jours, conjointement avec le pur sang anglo-arabe, dont nous nous entretiendrons, à améliorer nos races françaises ; mais, disons une fois pour toutes, que le type régénérateur restera le pur sang arabe, auquel il faudra avoir recours au cas d'abâtardissement de ses descendants.

173 sur 319

Les EPUB de la BnF

Ces EPUB constituent une transposition littérale de l'ouvrage original sous une forme numérique, sans transformation ni adaptation. En revanche, d'autres projets conduits par des partenaires de la BnF ont vocation à produire des livres numériques en EPUB de façon éditoriale (correction du texte original, nouvelle couverture, ajout d'un apparat critique...) Les livres numériques commercialisés dans le cadre du projet ReLIRE sont produits par FeniXX, tandis que ceux qui vont rejoindre l'offre bnfcollection.com de la filiale BnF-Partenariats sont fabriqués par Ligaran, un partenaire spécialisé dans le domaine de l'édition. De son côté, le Labex OBVIL produit dans le cadre de son activité de recherche des EPUB de textes littéraires classiques, dotés d'un apparat critique scientifique, qui ont vocation à être diffusés sous licence libre et intégrés dans Gallica.

Schéma numérique - mars 2016

06.B
Recherche et développement autour de l'océrisation

Ces dernières années, la BnF a fortement investi le champ de la recherche appliquée en reconnaissance optique de caractères (OCR). Il s'agit pour l'institution de communiquer à ses usagers des documents numérisés en mode texte de la meilleure qualité possible, une mission que peut seule satisfaire l'acquisition de savoir-faire et d'outils adéquats, conçus dans le cadre de travaux de R&D.

DE NOUVELLES PROCÉDURES DE CONTRÔLE QUALITÉ

La BnF a participé durant la période 2011-2014 à de nombreux projets de recherche et lancé un projet dans le cadre de son plan triennal en association avec l'université de Rouen (« Contrôle qualité des données OCR »). Les procédures ainsi développées sont adaptées au traitement de volumes considérables de fichiers numériques via des logiciels d'océrisation, afin d'être appliquées aux données livrées par les prestataires de numérisation.

MAÎTRISE DES STANDARDS

La BnF étant tributaire de formats techniques en constante évolution (METS, ALTO), la présence active de ses représentants aux comités éditoriaux des formats en question est nécessaire pour s'assurer de la richesse des informations qu'ils peuvent véhiculer.

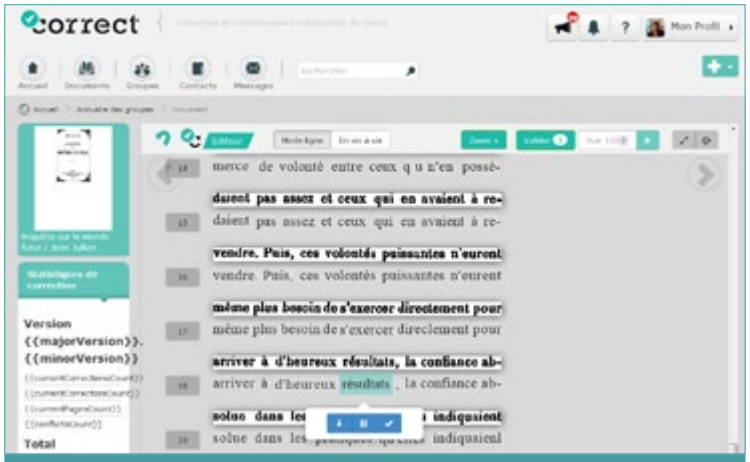
DE NOUVEAUX USAGES

Les traitements OCR doivent enfin être pensés à l'aune des besoins des usagers, qui évoluent avec les pratiques culturelles, et que les marchés d'océrisation doivent anticiper.
Le projet européen Europeana Newspapers (2012-2014) (voir axe 3 fiche N), destiné à numériser et valoriser les fonds documentaires de presse, a permis d'expérimenter la correction manuelle des contenus océrisés, la reconnaissance de la structure logique des fascicules (rubriquage, titrage, découpage de l'article), les possibilités de recherche en texte intégral.

UN CENTRE DE COMPÉTENCE EUROPÉEN

En tant que membre du Centre de compétence IMPACT (ONG promouvant la numérisation des textes patrimoniaux), la BnF sera pleinement associée à l'élaboration de nouveaux projets européens (notamment dans le cadre de l'appel à projets Horizon 2020), centrés sur les technologies OCR considérées dans leur périmètre le plus large : reconnaissance du texte, mais aussi correction automatique et enrichissement de l'information portée par le texte (reconnaissance d'entités nommées, extraction de concepts, etc.). À long terme, ces travaux aboutiront à des réalisations décisives, tels l'intégration d'un outil de correction automatique dans la chaîne de traitement des documents de presse numérisés, ou encore le développement d'un outil rendant possible le traitement rétrospectif des fonds de presse.

Il s'agit pour l'institution de communiquer à ses usagers des documents numérisés en mode texte de la meilleure qualité possible.



L'interface de correction de l'OCR de CORRECT

La plateforme CORRECT

Avec sa participation au projet de recherche FUI OZALID pour la conception de la plateforme CORRECT (Correction et enrichissement collaboratifs de texte), la BnF a expérimenté la production participative (*crowdsourcing*) en proposant à ses usagers d'améliorer eux-mêmes la qualité de l'OCR des documents numérisés de Gallica. Ce projet de R&D a fourni l'occasion d'explorer les possibilités d'enrichissement des documents pour en produire de nouvelles versions (livre électronique, édition adaptée aux non-voyants, etc.).

06.C Les moyens humains

Tout en ayant à cœur de maintenir l'expertise de ses personnels dans d'autres domaines (collections, métiers d'art…), la BnF souhaite les accompagner durablement dans la transition numérique. Du point de vue de la gestion prévisionnelle des effectifs, des emplois et des compétences (GPEEC), le numérique est en effet considéré comme un facteur d'évolution technologique ayant une incidence sur les emplois (émergence de nouveaux profils, évolution des activités et procédures…).

LE RÉFÉRENTIEL DES EMPLOIS ET DES COMPÉTENCES

À la BnF, la GPEEC repose sur un certain nombre d'outils, au premier rang desquels le *Référentiel des emplois et des compétences*, qui a été actualisé entre 2010 et 2012, notamment pour prendre en compte l'impact du numérique. Divers groupes de travail ont analysé, par catégorie d'emplois, les changements (observés ou prévisibles) introduits par les opérations de numérisation de masse ainsi que par les activités nouvelles (archivage du web, éditorialisation des contenus de Gallica, etc.).

MESURER L'IMPACT DU NUMÉRIQUE

Le principal défi en termes de pilotage des ressources humaines réside cependant dans la mise en place d'indicateurs permettant de mesurer l'effet qualitatif et quantitatif du numérique sur les emplois. En effet, s'il est facile d'identifier les agents exerçant des activités liées à la numérisation, il est en revanche plus difficile d'identifier les agents dont les activités sont totalement ou partiellement tournées vers le numérique, tant la diversité de ces activités est grande.

Le seul indicateur qui permet d'évaluer de manière objective la croissance du nombre d'agents impliqués dans des activités liées au numérique est le nombre de formations proposées et suivies par les agents dans ce domaine. Une enquête réalisée en 2013 a révélé qu'entre 2008 et 2013, le nombre de formations proposées dans le domaine du numérique est passé de 21 à 72. Jusqu'en 2007, moins d'une centaine d'agents étaient formés chaque année au numérique dans le cadre des activités cœur de métier (services au public, traitement des collections) ; en dehors de deux pics à plus de 600 en 2008 et 2013, ce nombre s'est ensuite stabilisé autour de 400 par an, soit environ 15 % de l'effectif de l'établissement.

ORH-ION : une instance de partage d'expérience et de réflexion

Créé au sein de la BnF en 2008, l'Observatoire de l'organisation et des ressources humaines sous l'impact opérationnel du numérique (ORH-ION) réunit régulièrement autour de séminaires thématiques les acteurs des programmes et activités numériques de la BnF. Affranchi de tout cadre hiérarchique et opérationnel, ORH-ION joue un rôle d'accélérateur dans l'émergence d'une culture partagée du numérique.

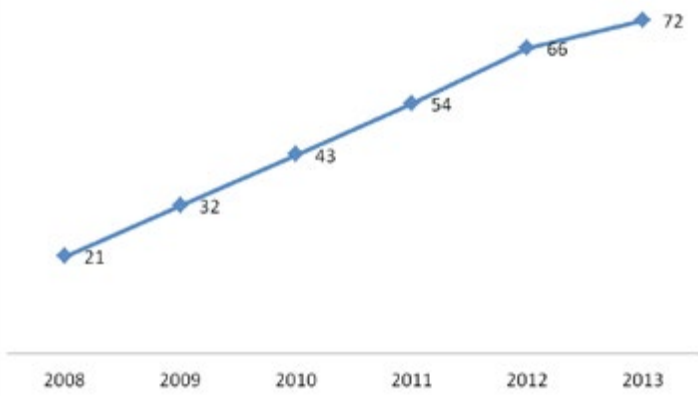
Parmi les questions de fond soulevées par les évolutions des métiers et des tâches qui ont pu être débattues dans ce cadre depuis 2010, on notera des retours d'expérience sur la médiation numérique ou encore la gestion des collections nées numériques. Le groupe a également conduit plusieurs enquêtes sectorielles : sur la fonction de signalement, sur le magasinage à l'heure du numérique ou encore sur l'évolution du métier de photographe.

DES MOYENS SPÉCIFIQUES

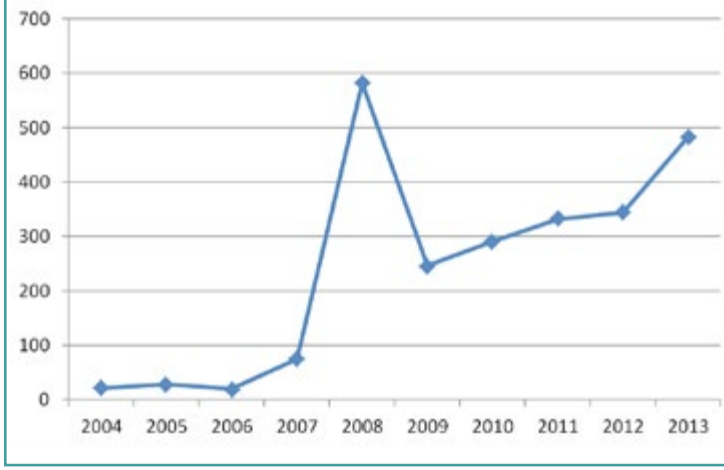
Certains projets ont toutefois fait l'objet d'un déploiement de ressources humaines dédiées, en particulier lors de la mise en place de nouveaux dispositifs. Le projet ReLIRE (voir axe 1 fiche G) bénéficie ainsi d'un financement fléché en ressources humaines de la part du CNL, à hauteur de près de 400 000 euros par an, soit 6 ETP. Quant au dépôt légal des livres numériques, sa mise en œuvre repose sur le redéploiement de sept ETP auparavant dédiés à la redistribution du 2e exemplaire du dépôt légal imprimé, à la faveur de la réforme qui a vu la réduction des dépôts à un exemplaire unique (2015).

Le principal défi en termes de pilotage des ressources humaines réside dans la mise en place d'indicateurs permettant de mesurer l'effet qualitatif et quantitatif du numérique sur les emplois.

Evolution 2008-2013 du nombre de stages proposés en lien avec le numérique



Evolution 2004-2013 du nombre de stagiaires formés dans le domaine du numérique : cœur de métier (services aux publics, traitement des collections)



06.D L'investissement dans le développement informatique

La capacité d'innovation de la BnF repose pour une grande part sur son aptitude à inventer et fabriquer les applications et logiciels capables de prendre en charge des fonctions nouvelles ou des volumétries dépassant largement les performances des progiciels du marché.

DES BRIQUES *OPEN SOURCE*

L'investissement consenti pour la réalisation de ces outils, souvent innovants d'un point de vue fonctionnel et technique, est parfois partagé avec des communautés grâce à des briques *open source*, non seulement pour des raisons financières, mais également par souci de pérennité, de respect des standards, de dynamique d'évolution et d'ouverture (c'est le cas en particulier des travaux concernant le dépôt légal du web : cf. axe 5 fiche F). Pour une part de cette offre logicielle, prise en charge sur des budgets alloués à la BnF par le CNL, il s'agit là d'un investissement national qui profite aux partenaires de l'établissement (établissements français, francophones, européens ; éditeurs ; acteurs du livre, etc.).

UN PLAN DE DÉVELOPPEMENT LOGICIEL

L'établissement s'est doté d'un « plan de développement logiciel » annuel listant l'ensemble des besoins logiciels tous domaines confondus, avec le détail des moyens financiers et humains nécessaires sur l'année à venir ainsi que les projections financières pour les deux années suivantes. Une fois ces investissements arbitrés, les équipes du département des systèmes d'information assurent la responsabilité de chaque projet dans toutes ses composantes (qualité, coût, délai, risque).

DES PORTEFEUILLES DE PROJETS

Des responsables de portefeuille gèrent un ensemble de projets cohérents. Ils accompagnent les équipes, jouent un rôle de conseil et d'expertise, assurent une veille technologique et prospective, et participent à la vie des projets de leur portefeuille (analyse des événements, arbitrages, information de la hiérarchie). Cette organisation, associée à l'utilisation de la méthode agile, favorise l'innovation et des résultats orientés vers l'utilisateur. Elle garantit une visibilité continue sur la progression des travaux grâce à des échanges constants entre les acteurs (informaticiens, prestataires, utilisateurs, partenaires, décideurs). Elle permet de revoir les priorités en cours d'année et d'intégrer des besoins nouveaux.
L'investissement dans le développement logiciel pour les projets numériques entre 2011 et 2014 a représenté plus de 16 millions d'euros, dont plus de 6 millions ont été apportés par le CNL.
Il a mobilisé, sur cette période, 33 à 36 ETP au sein du département des systèmes d'information.

L'investissement dans le développement logiciel pour les projets numériques entre 2011 et 2014 a représenté plus de 16 millions d'euros.

La méthode agile « Scrum »

Utilisée à la BnF depuis 2010, elle a joué un rôle déterminant dans la dynamique d'évolution des applications. Expérimentée, puis généralisée en 2013 aux différents projets de développement informatique des domaines numérique et bibliothéconomique, elle a favorisé la communication entre équipes et la compréhension mutuelle des besoins. Autorisant une planification plus souple des développements, une meilleure adéquation du produit au besoin, et permettant d'anticiper l'évolution des usages, la méthode a également fait émerger une nouvelle fonction, le chef de produit (*product owner*), chargé notamment de réunir et de hiérarchiser les besoins.

	2011	2012	2013	2014	Total
BnF	80 00 €	48 000 €	517 00 €	14 00 €	95 00 €
CNL	07 00 €	865 00 €	940 00 €	16 00 €	04 00 €
Autres	- €	- €	17 00 €	30 00 €	47 00 €
Total	88 01 €	35 01 €	476 01 €	4 34 01 €	16 04 00 €
ETP DSI	33,4	33,1	34,3	35,6	

Tableau récapitulatif de l'investissement lié au numérique dans le plan de développement logiciel (2011-2014)

06.E
Communication interne et accompagnement au changement

Le numérique innerve l'ensemble des métiers de la BnF. Devant cette évolution structurelle, un accompagnement au changement s'impose. La communication interne joue un rôle important en la matière.

LES « RENDEZ-VOUS DU NUMÉRIQUE »

En 2013, des séances de sensibilisation aux défis du numérique ont été mises en place, à raison de deux par an. Intitulées les « Rendez-vous du numérique », ces séances sont destinées à l'ensemble des agents. Leur objectif consiste à créer une culture commune du numérique et à mieux faire connaître les formations proposées par l'établissement. Le journal interne *Trajectoire* favorise lui aussi l'acculturation au numérique, à l'aide de focus sur un projet, une équipe, un service, une méthode, une innovation. Les réunions d'information « Midi de l'info » (mises en ligne sur l'intranet dans une version enrichie comme l'ensemble des séances d'information) permettent, quant à elles, de présenter des projets et des réalisations, ainsi que les grands axes de la stratégie numérique. Cet effort d'information et de pédagogie se nourrit de partages d'expérience avec d'autres établissements, comme la méthode de production participative de la Bibliothèque royale du Danemark, ou encore la transition numérique de l'Agence France-Presse.

UN RÉSEAU SOCIAL D'ORGANISATION

Au-delà de ces initiatives pédagogiques et informationnelles, la BnF s'adapte aux nouvelles pratiques des agents et aux outils qu'ils utilisent pour partager leurs savoirs et leurs expériences.
La mise en place d'un réseau social d'organisation (RSO) y contribuerait, en instituant une communauté d'intérêts permettant de gérer et suivre un projet, coconstruire des documents, créer un corpus de connaissances, recenser et rechercher des compétences, faire émerger des idées, etc. L'intégration d'un tel instrument collaboratif, en rupture avec les cultures d'organisation traditionnelles, appelle une évolution des pratiques managériales. Sa réussite implique un consensus autour de « règles du jeu » partagées, et l'implication du management en faveur du travail en réseau. Ainsi pourront être fluidifiés et dynamisés les flux d'information, favorisées la transversalité et l'efficacité, développée l'innovation participative, et enfin renforcée la motivation des « agents-acteurs ».

L'intégration d'un instrument collaboratif, en rupture avec les cultures d'organisation traditionnelles, appelle une évolution des pratiques managériales.

06.F La formation continue

Une fois la révolution du numérique concrètement intégrée dans ses missions, la BnF a vu émerger de nouvelles compétences, souvent développées de façon informelle et expérimentale par les experts. L'établissement a pris peu à peu conscience de cet accroissement collectif des savoirs et des savoir-faire, et a entrepris de l'accompagner par un effort de formation.

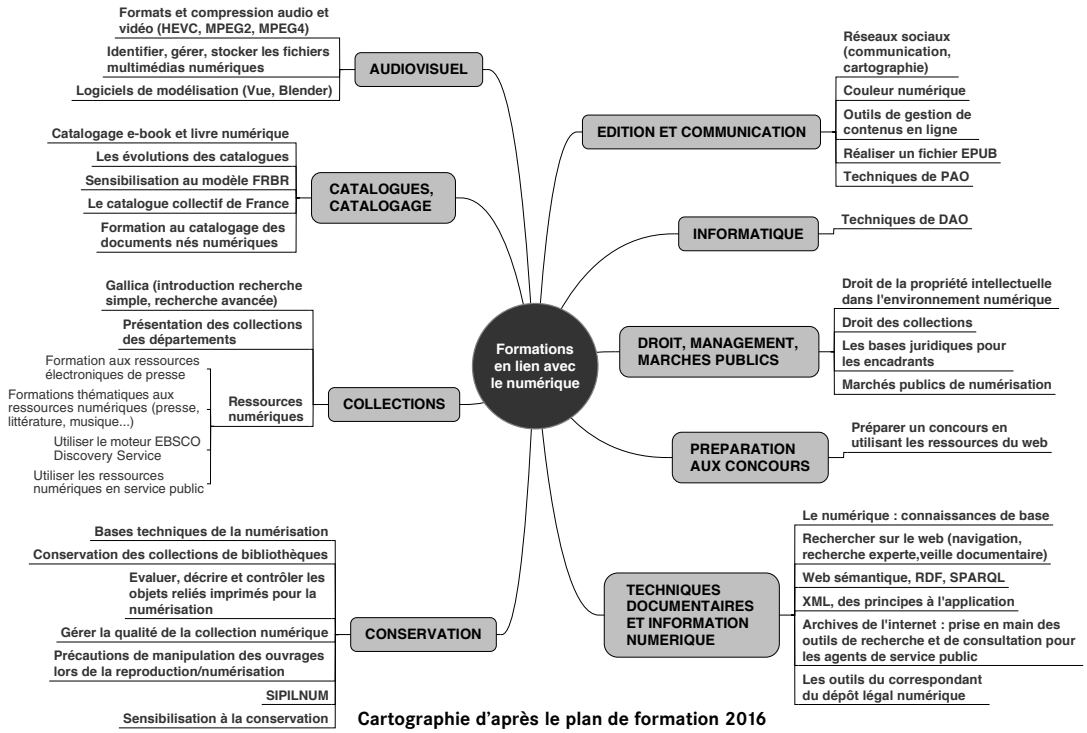
LE PLAN DE FORMATION

Il a progressivement intégré le numérique sous des formes variées : monitorat, tutorat, information, formation. Ces actions recouvrent plusieurs domaines du plan de formation de l'établissement : collections et ressources numériques, conservation, édition, techniques documentaires et information numérique… Des attentes en matière de gestion, de droit ou encore de marchés publics se sont exprimées : les personnels des corps des bibliothèques ne sont pas les seuls concernés par le développement du numérique.

DES PRATIQUES NOVATRICES

Notons que l'acquisition de compétences nouvelles ne peut s'effectuer dans le seul cadre de stages, les sujets les plus innovants ne figurant pas toujours au catalogue des organismes de formation établis. D'autres formes d'accompagnement, tels le tutorat, la participation à des ateliers en marge de conférences internationales spécialisées, ou encore le partage de la veille documentaire doivent dès lors être utilisées pour développer les connaissances et les savoir-faire dans les domaines les plus émergents.

La BnF encourage des formes d'accompagnement diversifiées, tels le tutorat, la participation à des ateliers en marge de conférences internationales spécialisées, ou encore le partage de la veille documentaire.



Cartographie d'après le plan de formation 2016

Un exemple de formation interne

Les services de la Bibliothèque se montrent créatifs pour garantir la transmission des expertises entre agents impliqués dans les projets numériques et ceux qui ne les perçoivent que de manière éloignée ou partielle. En accompagnement des projets d'innovation numérique qu'il pilote, le département du dépôt légal a lancé fin 2013 un cycle de formation interne sur le numérique. Cette initiative entendait favoriser l'acquisition d'une culture numérique commune par l'ensemble des agents du département. Sous la forme simple et interactive de l'atelier, la formation encourageait l'expression des participants, et ce d'autant plus aisément que les questions traitées se rapportaient à des situations vécues au quotidien. Le cycle s'est ouvert par deux séances plénières, complétées par des ateliers optionnels auxquels plus de 48 % des agents du département se sont inscrits. Cette initiative a de surcroît renforcé la transversalité au sein même du département.

Schéma numérique - mars 2016

06.G
La fouille de textes et de données (*text & data mining*)

Pour assurer l'accès du plus grand nombre aux collections, la BnF doit offrir plus qu'un simple accès aux contenus numérisés : elle doit permettre aux chercheurs d'utiliser les outils d'analyse qui permettent d'aller au-delà de ce que l'œil humain peut lire, pour établir de nouvelles connaissances. La fouille de textes et de données (ou TDM pour *text & data mining*) offre l'opportunité de voir émerger de nouvelles pistes : il ne s'agit plus de consulter un document pour son seul contenu, mais d'interroger la place qu'il occupe dans un corpus, questionner ses métadonnées, ses occurrences internes et externes.

DES PROJETS EN PARTENARIAT

Les données produites et conservées par la BnF, qu'il s'agisse de ses notices bibliographiques et d'autorité placées sous licence ouverte de l'État, des corpus numérisés et océrisés de Gallica, ou des archives de l'internet (voir axe 2 fiche A), ouvrent à cet égard d'immenses possibilités. Deux projets ont été lancés parallèlement : le premier, conduit avec le Laboratoire PACTE de Sciences Po Grenoble, consiste à analyser des corpus liés aux élections de 2012 et aux Jeux Olympiques de 2012 ; le second est mené auprès de Télécom ParisTech et de la Bibliothèque de documentation internationale contemporaine (BDIC), dans le cadre du Labex « Les Passés dans le Présent » (voir encadré). Ces deux initiatives ont pour objectif d'extraire des archives du web des métadonnées permettant de décrire, d'analyser, de caractériser et de cartographier les corpus.

À ces données de nature patrimoniale s'ajoutent celles que l'établissement crée et collecte au fil de son activité : traces de consultation des ressources de Gallica, parcours des lecteurs dans les salles, données administratives exploitées dans le cadre de l'Observatoire du dépôt légal… Autant de sources d'information dont les technologies des mégadonnées (*big data*) pourraient, après anonymisation, démultiplier la portée.

UN NOUVEAU SERVICE POUR LES CHERCHEURS

Demain, la BnF a vocation à développer cette activité, en proposant à ses chercheurs de bénéficier d'une infrastructure sécurisée leur permettant, suivant les données auxquelles ils s'intéressent et, en particulier, suivant le statut juridique des collections qu'ils étudient, de constituer leurs corpus de recherche et de les explorer soit avec leurs propres outils, soit avec ceux de la BnF.

Les technologies des mégadonnées (*big data*) pourraient, après anonymisation, démultiplier la portée des données créées et collectées par la BnF.

« Le devenir en ligne du patrimoine numérisé de la Grande Guerre »

Mené en collaboration avec la Bibliothèque de documentation internationale contemporaine (BDIC) et l'Institut des sciences sociales du politique (ISP), ce programme a été initié pour la période 2013-2016 dans le cadre du Labex « Les passés dans le présent ». Il a pour objectif de mieux comprendre la manière dont un corpus numérisé – celui de la guerre 14-18 – se diffuse à travers le web, y suscitant des échanges et des appropriations multiples. L'une des phases de l'étude s'est appuyée sur un corpus issu des archives de l'internet, rassemblé à l'occasion d'une collecte dédiée. Des outils de TDM ont été utilisés d'une part pour cartographier les liens entre les sites, d'autre part pour analyser les contenus du forum 14-18, l'un des principaux sites amateurs consacrés à la Grande Guerre.

06.H La numérisation en 3D

La nouvelle version de Gallica, développée en 2015, intègre la visualisation en 3D de globes conservés au département des cartes et plans.
Ce nouveau type d'objet numérique est le fruit d'un accord de mécénat de compétence pour la numérisation et la restitution en 3D d'un ensemble de 55 globes terrestres et célestes, conclu en juillet 2014 par la BnF et Dai Nippon Printing (DNP). DNP apporte à la BnF ses technologies de numérisation en trois dimensions, développées pour l'occasion. Avec cette réalisation, les processus de numérisation mis en œuvre par la BnF prennent une nouvelle dimension.

RESTITUTIONS NUMÉRIQUES EN VOLUME

Le département des cartes et plans conserve environ 120 globes anciens terrestres et célestes, dont plusieurs pièces majeures du patrimoine mondial. Cette collection exceptionnelle est un témoignage inestimable de l'évolution des connaissances géographiques et de la compréhension de l'univers connu dans l'Occident chrétien et dans le monde arabe, entre le XIe et la fin du XIXe siècle.

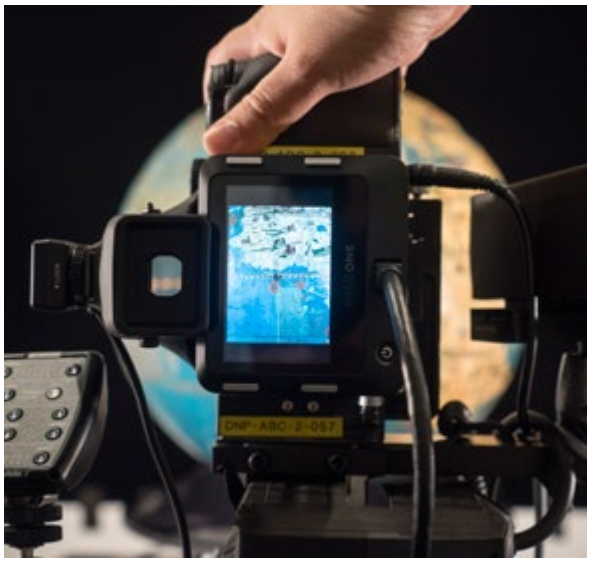
Le partenariat entre la BnF et DNP avait pour but de réaliser, à l'échelle d'une collection, des restitutions numériques en volume des sphères, qui soient conformes aux originaux, mobiles en tous sens de rotation, et en haute définition, afin de pouvoir zoomer avec fluidité et lisibilité sur toute la surface.

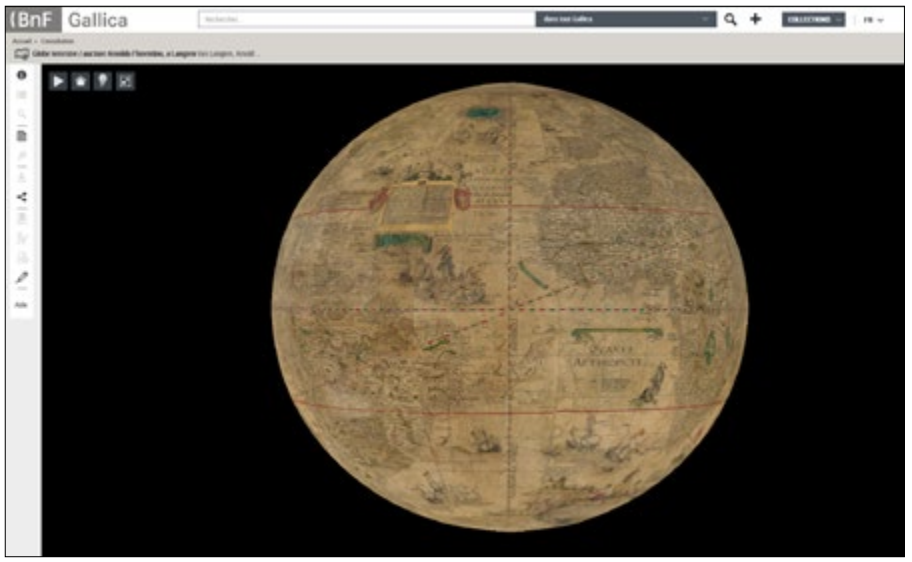
UNE REMARQUABLE COLLABORATION

Après la modélisation 3D des données numériques, réalisée au Japon par la société nipponne, la BnF archive les masters haute définition dans SPAR et diffuse sur Gallica une version plus légère. Grâce à cette remarquable collaboration franco-japonaise, ces objets pionniers sont désormais visibles dans la nouvelle version du visualiseur de la bibliothèque numérique.
En 2016, une exposition virtuelle en ligne, soutenue par la Fondation Total, proposera à un large public une exploration commentée de ces globes.
Des dispositifs innovants de présentation dans les espaces d'exposition de DNP et de la BnF sont également imaginés en vue de la présentation de cette réalisation à Tokyo et à Paris.

Les globes numérisés en 3D sont visibles dans la nouvelle version du visualiseur de la bibliothèque numérique.

Prise de vue photogrammétrique d'un globe terrestre dans les ateliers du département des cartes et plans.





06.1
Le projet de coopération avec le CEA

La BnF a engagé fin 2014 une démarche prospective visant à dessiner les contours d'une bibliothèque nationale pour la France de 2025. Impulsée à l'occasion d'un séminaire de l'encadrement animé par les équipes du CEA TECH, cette réflexion vise à imaginer le positionnement scientifique, culturel, éducatif et sociétal de la Bibliothèque dans dix ans. Quels seront ses publics ? Quels services devront être proposés, pour quels besoins et pratiques de recherche, de culture, de loisir ? Comment s'organiseront les missions de la BnF dans un environnement transformé par le numérique, et avec quels partenaires ?

L'écosystème partenarial est à redessiner entre opérateurs et partenaires avec lesquels la BnF coopère ou pourrait coopérer.

S'agissant spécifiquement de la place et des évolutions du numérique, la démarche innovante engagée par la BnF sera guidée par trois interrogations majeures.

UNE BIBLIOTHÈQUE NATIONALE POUR QUOI ET POUR QUI ?

Il s'agit d'imaginer les espaces physiques et virtuels d'une bibliothèque s'adaptant à de nouveaux usages et à une nouvelle économie de la connaissance. Dans un contexte de croissance des pratiques numériques et des usages en mobilité, de nouveaux outils et services numériques sont à inventer. L'expérience de l'usager doit être placée au centre de la culture professionnelle et repensée dans toutes ses dimensions.

UNE BIBLIOTHÈQUE NATIONALE AVEC QUI ?

L'écosystème partenarial est à redessiner entre opérateurs et partenaires avec lesquels la BnF coopère ou pourrait coopérer : pour la production de données bibliographiques, la constitution et la mise à la disposition de ressources, la production culturelle, la conservation patrimoniale, etc. Quelle place et quels services imaginer pour des institutions dépositaires, comme la BnF, de volumes importants de données et métadonnées « certifiées », dans une économie de la production et de la diffusion de la connaissance (recherche, enseignement…) profondément transformée dans ses méthodes ? De nouveaux environnements numériques de travail, espaces et réseaux, permettant la fouille de données ou la coconstruction de preuves de concept, peuvent voir le jour à la faveur d'opportunités comme l'installation de 1 000 start-ups innovantes, dès 2016, dans la Halle Freyssinet (près du site François-Mitterrand).

UNE BIBLIOTHÈQUE NATIONALE COMMENT ?

Les démarches mises en œuvre pour la conduite et l'accompagnement au changement (acculturation aux évolutions numériques, changement de posture vis-à-vis des publics, etc.) doivent elles-mêmes prendre de nouvelles formes, afin d'installer durablement la culture de l'innovation au sein de l'établissement. Une plus forte implication collective de tiers (personnels, usagers, parties prenantes, partenaires…) peut se faire jour via des approches d'innovation participative et en développant une expertise spécifique dans le domaine des méthodes de créativité.

La collaboration de la BnF avec l'Ideas Laboratory hébergé par le CEA se matérialisera ainsi en particulier par l'accueil d'une antenne du laboratoire, conçue comme un espace créatif ouvert aux membres de l'Ideas Laboratory et aux partenaires de la BnF, ainsi qu'aux personnels et publics désireux de participer à la conception de la Bibliothèque de demain et à l'expérimentation des projets les plus novateurs.

L'enjeu d'une veille organisée

La veille pratiquée à la BnF consiste à collecter, et éventuellement rediffuser, des informations jugées utiles ou stratégiques. Cette compétence nouvelle s'est développée en moins d'une décennie grâce aux outils du web 2.0 et aux réseaux sociaux. Elle a été prise en compte dès 2012 dans le *Référentiel des emplois et des compétences*. Pour optimiser la diffusion et la gestion des résultats de la veille, une cartographie de cette activité et de ses modalités de diffusion serait à entreprendre : enquête sur les besoins, leur taux de satisfaction ; identification des causes de l'absence de veille, priorisation éventuelle ; repérage des compétences internes en matière de formation à la veille. Dans le cadre posé par le *Schéma numérique des bibliothèques* de 2010, un tel outil permettrait d'envisager une réflexion sur la complémentarité entre les veilles professionnelles assurées en France. L'élaboration, au sein de la BnF, de scénarios d'organisation, de formation, de diffusion, d'évaluation et d'archivage de la veille pourrait se développer en lien avec l'activité d'innovation qui sera conduite dans le cadre de l'Ideas Laboratory.

Crédits photographiques : Eric Sempé / BnF (Couverture, p. 9, 21, 42-43, 82), Paolo Verzone-VU' / BnF (p. 14-15, 84-85), David Paul Carr / BnF (p. 22, 26, 49, 62, 65, 116), Luc Verrier / BnF (p. 23), Alain Goustard / BnF (p. 30-31, 73), Patrick Bramoullé / BnF (p. 37), Thierry Ardouin / Tendance Floue / BnF (p. 39), Pascal Lafay / BnF (p. 46), Jennifer Boche (p. 59), David Goehring / Flickr (p. 63), Catherine Floc'hlay (p. 82), Béatrice Lucchese / BnF (p. 92), Brigitte Pougeoise / Fonds Brigitte Pougeoise / BnF, (p. 100), DR, BnF (p. 101-103), DR.

Conseil et création : Kstories

BnF | Bibliothèque nationale de France

François-Mitterrand
Quai François Mauriac, Paris 13e _ 33 |0|1 53 79 53 79

Richelieu
5, rue Vivienne et 2, rue Louvois, Paris 2e _ 33 |0|1 53 79 53 79

Bibliothèque de l'Arsenal
1, rue de Sully, Paris 4e _ 33 |0|1 53 79 39 39

Bibliothèque-musée de l'Opéra
Place de l'Opéra, Paris 9e _ 33 |0|1 53 79 37 40

Maison Jean Vilar
8, rue de Mons, 84000 Avignon _ 33 |0|4 90 86 59 64

Centre technique (CTBnF)
14, avenue Gutenberg, 77607 Bussy-Saint-Georges _ 33 |0|1 53 79 38 44

Centre de conservation Joël Le Theule
Le Château, 72300 Sablé-sur-Sarthe _ 33 |0|2 43 95 19 92

www.bnf.fr

© BnF | direction des services et des réseaux, délégation à la communication, mars 2016