

Le numérique à la BnF : un objectif global

Par Laurence ENGEL

Présidente de la Bibliothèque nationale de France

La Bibliothèque nationale de France est depuis plus de 25 ans un opérateur numérique. Elle en maîtrise toutes les facettes et en exploite les possibilités pour réaliser l'ensemble de ses missions – conservation, numérisation, entrée des productions nées numériques, gestion patrimoniale du Web, catalogage, valorisation... –, la conduisant à formaliser sa stratégie dans une approche globale et à la réactualiser régulièrement. Acteur transversal, à la croisée de toutes les industries culturelles et créatives, et habituée à partager ses outils comme son patrimoine, elle tente de sensibiliser ses partenaires à une vision « totale » d'une technique qu'il s'agit de maîtriser pour ne pas y être soumis. Elle est aussi décidée, au-delà de son orientation résolument technologique, à rappeler que le monde physique est notre premier destin.

Voilà plus de vingt ans que la Bibliothèque nationale de France a opéré sa mue numérique. Cela lui donne – déjà –, parmi les institutions culturelles, les allures d'un vieil opérateur aguerri à cette technologie ! Mais aussi l'occasion de rappeler la proximité ontologique entre le monde des bibliothèques et celui du Web (qui partagent les notions de fichier, de référencement, de catalogage, de signalement...), proximité qui explique sans doute la capacité qu'elles ont eue à se projeter dans ce nouveau monde sans perdre quoi que ce soit de leur identité. Et de rappeler enfin l'énergie qu'il a pourtant fallu déployer pour convaincre – parfois au sein même de nos institutions, souvent parmi nos interlocuteurs, notamment ceux de la sphère politique et administrative – que cette mue était à notre portée, et que nous n'avions nul besoin de nous livrer pieds et poings liés aux grands opérateurs américains pour nous mettre à niveau. Ces trois éléments stratégiques – l'anticipation, le respect des missions et des identités, et le volontarisme politique – continuent d'être déterminants pour assurer de manière autonome, souveraine et efficace le déploiement d'une action dans le monde numérisé qu'est devenu le nôtre.

Faire le point sur la stratégie numérique de la BnF suppose donc, pour commencer, d'en retracer l'histoire et le contenu. Mais aussi de rappeler la situation singulière de la Bibliothèque, au confluent de toutes les industries culturelles et créatives (ICC). Cette stratégie est de ce fait destinée à adapter la production de la Bibliothèque aux technologies les plus contemporaines – laquelle est, à ce titre, confrontée aux mêmes questions que se pose tout opérateur. Mais elle est aussi de penser ces évolutions en intégrant dans la réflexion l'amont et l'aval de sa production, là où se situent précisément les ICC, à la fois fournisseuses de biens, à travers l'obligation du dépôt légal, et bénéficiaires de l'action de la Bibliothèque.

Mais commençons par évoquer le Schéma numérique de la BnF qui rend compte de sa stratégie.

Il est le produit, depuis sa première version, à la fin des années 2000, d'une histoire du numérique à la BnF, que l'on peut résumer en trois étapes principales. La première est nettement symbolique. Il s'agissait de réagir à la proposition de Google de devenir l'acteur exclusif de la numérisation du patrimoine livresque, en France comme ailleurs. La tentation fut grande, au début des années 1990, de s'engouffrer dans une trajectoire qui pouvait paraître à bien des égards économe – Google s'occuperait de numériser gratuitement les ouvrages, de fournir aux institutions propriétaires de l'original une copie de cette numérisation et assurerait pour un temps, selon un contrat d'exploitation exclusive, la diffusion de cette bibliothèque numérique. Le président de la BnF d'alors, Jean-Noël Jeanneney, y vit une main basse sur le patrimoine français et la perte de notre souveraineté sur l'un des symboles de la Nation. Il milita donc contre l'acceptation de ce contrat et obtint gain de cause : ce fut l'acte 1 de la vie numérique de la BnF. Et Jean-Noël Jeanneney avait raison : ceux qui renoncèrent alors à s'occuper par et pour eux-mêmes de leur avenir numérique ont perdu un temps précieux et n'occupent plus aujourd'hui la place qui était alors la leur dans l'économie des savoirs. L'ambition – considérable sur le plan politique – était toutefois encore modeste sur le plan pratique, puisqu'elle prenait appui sur la notion de vitrine : il ne s'agissait pas encore de dupliquer la bibliothèque physique dans le monde numérique, mais simplement de l'illustrer en y donnant à voir quelques joyaux. L'étape suivante, en réalité la suite logique de ce premier geste, a donc consisté, après notre entrée dans le monde numérique, à en suivre l'évolution, à en prendre tous les plis – et notamment celui de la quantité, facteur de visibilité. Ce fut le début de l'époque industrielle, celle de la numérisation de masse pour atteindre la « taille critique ». Acte 2, tout aussi stratégique que le premier : pour être véritablement acteur numérique, il fallait passer de la vitrine au magasin tout entier. Ce qui a conduit à « numériser » progressivement l'ensemble des fonctions de la Bibliothèque : faire le lien entre la reproduction des

documents eux-mêmes et le processus déjà engagé de numérisation des catalogues (leur rétro-conversion pour en assurer la gestion et la visibilité en ligne) ; étendre la réflexion de la numérisation au stockage ; propulser la politique de valorisation à l'ère du Web. Mais cette deuxième étape était encore celle d'une construction parcellisée à laquelle manquait l'expression d'une vision d'ensemble, d'une stratégie. Ce fut le troisième temps de l'histoire numérique de la Bibliothèque : celui qui depuis plus de dix ans maintenant lui permet de mettre en avant un programme global à travers son Schéma numérique, dont nous avons finalisé, juste au moment de la crise sanitaire, la troisième version.

Cette édition prend la forme presque onirique d'une carte (géographique) doublée d'un jeu de cartes (permettant de se repérer dans la masse des projets engagés). Ce choix collectif et inventif résulte de l'état des lieux fait au préalable : après une première version représentant nos projets comme une carte heuristique organisée selon les missions de la Bibliothèque, en 2009, carte qui s'est tellement enrichie au fur et à mesure du développement des projets qu'elle en était devenue presque illisible ; après une deuxième version, préparée en 2016 sous la forme d'un document rédigé, classique en somme, et qui permettait d'explicitier chacune de nos démarches numériques, nous sommes donc revenus à la forme visuelle. Elle permet d'insister sur la dimension panoramique de nos projets, sur la présence du numérique dans l'ensemble des territoires que nous occupons, sur la complexité des démarches engagées, mais aussi sur ce qu'elles produisent : un paysage ouvert dans lequel chacun peut naviguer, choisir d'aller vers des sites remarquables et très fréquentés – nous y avons inscrit les « hauts lieux du numérique à la BnF », tels que le dépôt légal numérique, l'environnement de travail, Gallica, le catalogue général ou encore les outils de fouille –, ou, au contraire, s'aventurer dans des contrées encore peu explorées... Pour en résumer le dispositif, nous dirons que cette carte est organisée en huit régions : les collections numériques, les données, les infrastructures, le traitement des collections, le signalement et le catalogue, les technologies, le partage des ressources et, enfin, les projections (vers d'autres terres). Nous dirons ensuite qu'il permet d'identifier cinq axes stratégiques : abolir la fracture numérique ; étudier, observer, devancer les usages ; partager les infrastructures, les compétences et les expertises ; urbaniser les applications ; et faire évoluer les métiers. On voit à travers eux que le numérique concerne donc autant la gestion de l'établissement et la relation éthique avec nos usagers, que notre volonté de placer les publics au cœur de la mise en œuvre de nos missions, en passant par la coopération avec nos partenaires et la gestion prévisionnelle des compétences. Nous dirons enfin que la logique de ce schéma est celle des réseaux, qui permettent de mettre en relation et de montrer que ce que nous faisons ne vaut qu'au regard de ces relations : relations avec le public et avec les partenaires, relations avec les machines (automatisation, traitements par flux, interfaces de médiation), relations avec les technologies⁽¹⁾.

⁽¹⁾ Le Schéma numérique de la BnF : <https://www.bnf.fr/fr/mediatheque/schema-numerique-2020-de-la-bnf>

Pour compléter cette présentation rapide, je voudrais insister sur deux points.

Le premier est celui de la complétude de la démarche. Le numérique n'est pas seulement pour nous une surcouche visant à ajuster notre action aux exigences du monde moderne, il est aussi véritablement l'utilisation systémique d'un outil industriel au bénéfice de toutes nos missions. Si la BnF n'est pas *exclusivement* numérique, elle est *totale*ment numérique.

Cela vaut bien sûr pour la valorisation du patrimoine. Il peut s'agir d'ajouter des contenus à la présentation d'une œuvre, de pénétrer en elle (l'effet admirable du zoom), de la relier à son passé, à son environnement ou à sa postérité, de la reproduire et de la transporter loin de son lieu de conservation et d'aller ainsi à la rencontre de bien des nouveaux publics. Ce sont des expositions en ligne – bientôt répertoriées dans un véritable service en ligne de produits éditoriaux – ; des applications – par exemple, Gallicadabra qui propose pour les enfants des lectures et des jeux, ou encore BDnF pour tous ceux qui veulent s'essayer à la production de bandes dessinées en utilisant les ressources iconographiques de la Bibliothèque – ; des éditions numériques en ligne ou *in situ* – nous sommes ainsi actuellement en pleine phase de production pour accompagner sous format numérique l'ouverture du Musée de la BnF à Richelieu – ; des *podcasts* – pour créer des rendez-vous avec les collections comme le propose le programme « La BnF dans mon salon » ou pour accompagner RetroNews, le média en ligne consacré aux collections de presse... et, bien sûr, depuis plus de vingt ans, Gallica, la plus importante bibliothèque numérique patrimoniale au monde – près de 9 millions de documents –, un outil de recherche certes, mais aussi une porte d'entrée éditorialisée vers des collections inouïes. Et une bibliothèque qui se partage et se diffracte, pour des partenaires qui se comptent en centaines et font parfois le choix – c'est le programme « Gallica marque blanche » – d'utiliser notre technologie, éprouvée, pour donner forme à leurs propres projets documentaires ; qui se déploie aussi partout dans le monde, non pas seulement mécaniquement à travers le Web, mais aussi volontairement à travers la mise en commun virtuelle de collections qui racontent une histoire partagée – c'est la collection de bibliothèques numériques « Patrimoines partagés » et, plus largement, le résultat d'une politique active de coopération scientifique. C'est aussi le BnF-DataLab, le dernier né des services numériques de la Bibliothèque, un espace physique celui-là – un vrai laboratoire destiné à accueillir des équipes de recherche –, mais consacré aux données et à leur exploitation. Au total, le résultat s'exprime, de manière peut être trop abrupte mais néanmoins significative, dans les chiffres de fréquentation : si nos sites physiques drainent quelques 1,4 M de visites (en tout cas avant la crise sanitaire !), toutes activités confondues (salles de lecture, expositions, conférences et spectacles), le nombre de ces visites s'élève à 19 M sur Gallica, à 7 M sur RetroNews, et à environ 4 M pour les expositions numériques, pour ne citer que les propositions virtuelles qui attirent le plus grand public. L'effet du numérique est évidemment massif. Il permet aussi de mieux comprendre nos publics, de

mieux en appréhender les usages ; et c'est pour nous essentiel.

Mais au-delà, ce sont bien tous les outils de production et tous les métiers de la Bibliothèque qui sont impactés. Pour n'évoquer que les grandes phases du processus de patrimonialisation, il s'agit non seulement de numériser les collections physiques, mais aussi d'accueillir les collections nées numériques, de savoir en assurer la visibilité à travers le catalogage numérique et la maîtrise du référencement. Il faut aussi savoir conserver de manière pérenne ces collections numériques ou numérisées, tout comme il est essentiel d'utiliser des techniques prédictives pour ajuster nos politiques de conservation physique, de faciliter l'usage des collections par la mise en relation des données, d'en démultiplier l'exploitation en permettant l'accès à des outils et à des formats de recherche divers (avec toutes les problématiques liées aujourd'hui à la fouille). Les chantiers ne manquent pas – et l'on peut citer : côté conservation, SPAR (Système de préservation et d'archivage réparti), que l'on adapte aujourd'hui pour pouvoir développer notre activité de tiers archivage ; côté chaînes d'entrée des collections, MISAOA (Mutualisation et innovation pour la sauvegarde et l'accès aux œuvres audiovisuelles), développé en partenariat avec le CNC ; ou encore, côté catalogage, NOEMI (Nouer les œuvres, expressions, manifestations et *items*). Sans oublier, bien sûr, notre nouvelle feuille de route consacrée à l'intelligence artificielle et qui, même si y sont définis des projets précis, selon un calendrier séquencé, vise une nouvelle fois à inscrire une technologie dans l'ensemble des fonctions de la Bibliothèque pour mieux les appréhender.

Le second point sur lequel il me paraît important d'insister consiste à prendre la mesure des bénéfices que cette stratégie numérique peut apporter à l'ensemble des filières des ICC.

L'exercice du dépôt légal, dont il faut rappeler qu'il s'applique certes au livre, et ce depuis 1537, mais aussi à toutes les autres formes de produits éditoriaux – les partitions musicales, les estampes, les cartes, la presse, la photographie, l'édition phonographique, la vidéo, le jeu vidéo... , jusqu'au Web – entrés dans son champs au fur et à mesure de leur venue au monde, lui offre en effet un positionnement rare, totalement transversal. Mais cela lui impose aussi des devoirs. Les ICC nous apportent notre « matière première » et en retour nous assurons pour elles la conservation pérenne de leurs productions, rendant possibles pour l'avenir toutes les formes de valorisation. Cela vaut pour les œuvres mais aussi pour les données : les opérations de catalogage se font ainsi de plus en plus souvent à partir des données produites par les éditeurs eux-mêmes, lesquelles sont « dérivées », puis normalisées et enrichies par la BnF. Et, en retour, les éditeurs disposent librement de cette production de données, raffinées par la BnF, et de la possibilité de les relier les unes aux autres. À ce titre, le développement d'outils tels que *data.bnf* – qui permet d'interconnecter des collections issues d'institutions les plus diverses – ou *api.bnf* – qui permet d'associer les collections et les données de la BnF à des applications ou à d'autres univers grâce à des API – constitue un

apport essentiel à la vie en ligne des œuvres. La notion d'écosystème s'avère en la matière de plus en plus pertinente.

Mais la BnF doit sans doute mieux faire connaître ce rôle qui lui incombe. L'État met en avant à juste titre la nécessité d'assurer pour les œuvres françaises (ou francophones, et même européennes) une bonne « découvrabilité ». L'économie du Web repose sur la capacité à y être vu. Mais qui sait que la BnF, acteur reconnu parmi les bibliothèques nationales, est un catalyseur déjà opérationnel de cette attente ? À travers, par exemple, les organes de normalisation des données – ces lieux historiques du travail en commun des bibliothèques qui attirent d'ailleurs l'attention des Gafam : et ce parce que la normalisation détermine l'efficacité du Web, et plus précisément la capacité des « autorités » à être bien repérées, « découvrables » et suivies dans leur parcours. Et les bibliothèques nationales jouent, dans cette quête d'efficacité, un rôle historique et toujours pertinent. Là se construit une nouvelle valeur patrimoniale. C'est la raison pour laquelle, s'inscrivant dans la logique des États généraux des industries culturelles et créatives lancés par le ministère de la Culture, la BnF s'est engagée dans l'élaboration de projets communs à toutes les ICC. L'un de ces projets consiste par exemple, comme cela a déjà été fait dans les secteurs de l'agriculture, de l'agroalimentaire ou de la santé, à créer ensemble une plateforme d'échanges de données, une étape indispensable à l'exploitation de cette valeur. Intitulé « Alliance culture *data* », ce projet doit permettre aux filières culturelles, qui tardent à s'organiser pour éviter que la valeur des données qu'elles produisent ne soit exploitée par d'autres que par elles, de se saisir de ce sujet. N'est-ce pas là une nouvelle expression de cette volonté de préserver notre autonomie, notre souveraineté ? De rendre les œuvres visibles, accessibles ? Et de préserver les intérêts – les droits – de ceux qui les ont créées ? Ces projets sont à notre portée. Et l'espoir de la BnF est d'arriver à stimuler une réflexion commune, et d'aller encore au-delà de notre propre stratégie qui nous amène à toujours partager nos collections, nos données, nos outils.

Le numérique à la BnF ressemble donc à ce que la BnF est en elle-même : c'est le ramener à ce qu'il est, une technique dont il faut se saisir pour pouvoir la maîtriser. Ce faisant, la BnF ne fait preuve d'aucun excès d'enthousiasme – pas de dévotion, pas de soumission. Le numérique est tout autant une question qu'une solution. Il doit être interrogé. On ne peut poursuivre aujourd'hui dans la voie ouverte par les premiers actes de rétroconversion des catalogues, les premières numérisations, les premières médiations numériques... , sans poser le principe de la sobriété. Surtout, on ne peut poursuivre dans les démarches engagées sans rappeler que toutes les formes numériques de diffusion ne remplaceront jamais le rapport à l'œuvre dans sa matérialité ou le confort qu'offre une salle de lecture. Ce n'est pas là une conviction, mais l'affirmation d'une volonté et la prédisposition à un combat, car on ne saurait se satisfaire des solutions que les confinements successifs nous ont incités à produire.