



HAL
open science

Le numérique en bibliothèque : naissance d'un patrimoine : l'exemple de la Bibliothèque nationale de France (1997-2019)

Emmanuelle Bermes

► **To cite this version:**

Emmanuelle Bermes. Le numérique en bibliothèque : naissance d'un patrimoine : l'exemple de la Bibliothèque nationale de France (1997-2019). Histoire. Paris, Ecole nationale des chartes, 2020. Français. NNT : 2020ENCP0001 . tel-02475991

HAL Id: tel-02475991

<https://theses.hal.science/tel-02475991>

Submitted on 12 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE NATIONALE DES CHARTES – PSL

École doctorale 472

Emmanuelle Bermès

Archiviste paléographe

Diplômée d'études approfondies

Conservatrice générale des bibliothèques

LE NUMÉRIQUE EN BIBLIOTHÈQUE :
NAISSANCE D'UN PATRIMOINE

L'EXEMPLE DE LA BIBLIOTHÈQUE NATIONALE DE FRANCE
(1997-2019)

TOME PREMIER

Mémoire pour l'obtention du doctorat sur travaux
2020

Remerciements

Mes remerciements vont en premier lieu à Christophe Gauthier pour m’ avoir encouragée à me lancer dans l’ aventure du doctorat sur travaux, et pour m’ avoir soutenue et guidée au long de cette année, me mettant sur la voie d’ une réflexion nouvelle sur la naissance du patrimoine numérique. La vie est faite de rencontres : c’ est à l’ École des chartes que je fis celle, décisive, qui devait me mettre sur le chemin du numérique ; c’ est dans la même maison – quoiqu’ en d’ autres murs – que je devais trouver le moyen de tisser de nouveaux liens entre l’ estampe du XVII^e siècle et les octets circulant sur les réseaux.

Depuis 2003, mon parcours à la BnF a lui aussi été jalonné par de très nombreuses rencontres, aussi bien au sein de l’ établissement qu’ en France et à l’ international. Ces professionnels brillants, inspirants et motivés, au premier rang d’ entre eux Catherine Lupovici, ont contribué à forger collectivement une culture numérique à laquelle ces pages ne sont qu’ une contribution. Je remercie tout particulièrement les co-auteurs des différents articles recueillis dans le volume de travaux. Ces dernières années, le projet Corpus m’ a amenée à croiser ou recroiser les pas de différentes personnes, investies dans la recherche en histoire, en informatique, en sociologie ou dans d’ autres domaines, qui ont contribué à raviver mon goût pour l’ approche scientifique et à qui la présente entreprise doit beaucoup. Ma reconnaissance s’ adresse également aux collègues et anciens collègues de la BnF, pour certains retraités ou ayant vogué vers d’ autres rives professionnelles, et aux membres du comité d’ histoire de la BnF pour leurs encouragements et leur curiosité à l’ égard du travail de synthèse que je présente aujourd’ hui.

Il est enfin parmi mes proches, collaborateurs quotidiens, famille et amis, ceux qui ont partagé intimement chaque étape de l’ élaboration de cette réflexion, ceux qui ont exercé leur œil scrutateur sur ces pages, ceux qui m’ ont accompagnée dans les heures difficiles, ceux qui me stimulent chaque jour pour pousser encore plus loin ces réflexions. Qu’ ils soient remerciés pour leur soutien.

Introduction

Traiter aujourd'hui de la transformation numérique d'une institution comme la Bibliothèque nationale de France, telle qu'elle s'accomplit depuis une vingtaine d'années, est une gageure à bien des égards. L'institution est vaste, ses projets sont nombreux, ses collections d'une infinie diversité : l'impact de l'informatique, puis du numérique sur de nombreuses facettes de son activité nécessiterait une étude approfondie et de longs développements pour être couvert de manière exhaustive. Malgré une position de précurseur dans ce domaine, l'établissement continue de s'interroger régulièrement sur cette question, en son sein dans le cadre de démarches stratégiques¹ comme sur la scène internationale². Mais il ne s'agit pas seulement d'un enjeu professionnel, ou spécifique au domaine des bibliothèques : des industries culturelles aux rouages de la démocratie, de l'élaboration des savoirs aux mécanismes de l'économie, c'est la société tout entière qui, depuis le tournant du siècle, se transforme avec l'émergence d'une culture numérique qui accélère, globalise et bouleverse le monde. C'est pourquoi, en s'emparant du numérique comme questionnement fondateur sur ses missions et son identité, la Bibliothèque nationale joue pleinement son rôle d'opérateur culturel au service de la nation.

Au regard de tels enjeux, l'essai que nous présentons ici en vue de l'obtention du doctorat sur travaux ne peut proposer qu'une approche partielle et synthétique du sujet. Ayant pour objectif de contextualiser les textes publiés au cours des quinze premières années de notre

1 Par exemple, les différentes versions du *Schéma numérique*, l'élaboration des contrats d'objectifs et de performance...

2 On peut ainsi mentionner la conférence donnée par Laurence Engel, présidente de la BnF, à Stanford le 29 octobre 2018, ou encore les interventions des présidents et présidentes de la BnF au sein de la CDNL (Conférence des directeurs de bibliothèques nationales) lors du congrès de l'IFLA (International Federation of Library Associations). Par exemple : Bruno Racine, « How the national libraries will succeed with their digital strategies ». Présentation pour la réunion annuelle de la CDNL, Cape Town, 2015. En ligne : http://www.cdnl.info/sites/default/files/docs/2015_BnF_ADM_2015.pdf.

carrière de conservatrice des bibliothèques³, il impose inévitablement un exercice d'ego-histoire, interrogeant les faits et événements au prisme d'un point de vue particulier. Il ne s'agit pas, en outre, de répéter ni même de résumer le contenu de ces écrits précédents : le défi est d'apporter, dans un format resserré, une lumière supplémentaire à des considérations qui ont été formulées, au cours des années, dans un contexte de constante et profonde mutation.

Comment peut-on donc traiter de la transformation numérique d'une organisation quand on l'a vécue personnellement, alors même que cette transformation n'est pas achevée ? Disposons-nous du recul nécessaire pour en proposer une lecture historique ? Si nous avons en effet choisi de développer une large partie de notre propos selon une approche chronologique, c'est parce que le risque de perte de mémoire est réel pour l'institution : les changements sont rapides, les compétences, les personnes, les organisations évoluent vite.

L'histoire de la bibliothèque royale, impériale, nationale, plus récemment devenue « de France », si riche et passionnante, mobilise de nombreux esprits et suscite la publication de fréquents travaux⁴. Les événements de ces vingt dernières années ont pleinement vocation à s'y inscrire. Notre parcours professionnel de 2003 à 2019 a posé sous divers angles la question de l'impact du numérique sur le patrimoine, conduisant à observer son évolution et à tenter d'en comprendre les implications à la fois humaines et documentaires – une démarche indéniablement scientifique du point de vue des sciences de l'information et de l'histoire culturelle. La conjonction de ce parcours et de la profonde transformation de la BnF au cours de la même période nous semble dès lors pouvoir être vue comme une opportunité : celle d'un exercice, sinon d'histoire, du moins de mémoire.

Une mémoire qui s'attachera, parmi bien d'autres approches qui auraient pu être envisagées, à la question de l'émergence, depuis la fin du vingtième siècle, d'un nouveau patrimoine. Nous l'avons rassemblé sous le terme de « numérique », avec l'ambition de donner tout son sens, au fil des pages qui suivent, à cette expression imprécise.

Ancienne élève de l'École des chartes, nous avons consacré nos années d'initiation à la recherche scientifique à un tout autre type de patrimoine : l'imagerie populaire religieuse, à

3 Ces textes sont rassemblés dans le second volume ; afin de faciliter la navigation entre les deux tomes, les références des articles concernés sont indiquées en note de bas de page par deux astérisques (**). On les retrouve dans le volume de travaux classés par date et titre de publication.

4 Voir le site Internet du Comité d'histoire de la BnF : <http://comitehistoire.bnf.fr/>.

travers l'étude d'un couvent parisien au XVII^e siècle⁵. Il ne faut cependant pas voir, dans notre intérêt – survenu par la suite – pour le numérique, une rupture de trajectoire : d'un patrimoine à l'autre, on peut identifier une continuité. Le corpus d'une centaine d'estampes que nous avons alors rassemblé était en effet composé d'objets qui n'avaient pas été conçus, à l'origine, pour être conservés. Ces images de piètre qualité que l'on distribuait à l'église, pliait au fond d'une poche, glissait entre les pages d'un livre de piété ou placardait au mur d'une chambre frustre, n'ont été préservées à travers les siècles dans les collections des bibliothèques que par le fait d'un processus de patrimonialisation, relevant soit d'individus ou de communautés notamment religieuses, soit de dispositifs institutionnels comme le dépôt légal. Collectées, décrites et conservées, ces images de peu de prix se retrouvent désormais classées dans les magasins du département des estampes aux côtés des magnifiques eaux-fortes de Rembrandt et des sublimes burins de Claude Mellan, préservées avec le même soin et le même enjeu de transmission aux générations futures.

Notre environnement documentaire numérique actuel est d'un usage tout aussi quotidien et volatile, qu'il s'agisse des articles scientifiques que nous lisons en ligne, des billets de blog que nous écrivons, des plateformes qui nous fournissent musique et films illimités, des cartouches de jeux vidéo que nous glissons dans nos consoles, de l'article de presse que nous lisons sur notre écran de téléphone dans le métro, de l'enluminure médiévale dont nous partageons l'image sur les réseaux sociaux. Pourtant, la question de sa préservation et de sa patrimonialisation s'inscrit naturellement en continuité avec celle des objets de toute nature que la BnF conserve au sein de ses collections, dont l'imagerie populaire moderne fait partie. C'est donc en nous focalisant sur la problématique de ce patrimoine numérique émergent, à travers l'étude de la constitution des collections de la bibliothèque, que nous proposons d'aborder la transformation qui a marqué la BnF au cours des vingt dernières années.

Le choix d'une date de départ pour une telle analyse présente inévitablement un caractère arbitraire. Celle de 2003, qui correspond au début de notre parcours professionnel, ne nous semblait pas significative au regard du sujet à traiter : il nous fallait donc couvrir les années antérieures. Mais jusqu'où remonter ? La fusion administrative entre la Bibliothèque nationale et l'Établissement Public Bibliothèque de France (EPBF) en 1994 est davantage un

5 Emmanuelle Bermès, *Le couvent des Mathurins de Paris et l'estampe au XVII^e siècle*. Thèse pour le diplôme d'archiviste paléographe, 2001.

jalon organisationnel qu'un pivot essentiel du point de vue de la stratégie numérique. Celle-ci trouve ses racines dès la préfiguration du nouvel établissement et même, dès la formulation en 1988 de l'idée initiale de Jacques Attali d'une bibliothèque « d'un type entièrement nouveau » : cette expression, qui porte en son sein de manière implicite l'idée de transformation numérique, servira de fil rouge à notre propos. Elle se concrétise dès 1990 par un premier projet de numérisation, accompagné d'une réflexion approfondie sur la mise en place d'un dispositif de consultation très avancé pour les chercheurs : le Poste de Lecture Assisté par Ordinateur (PLAO). Pourtant, l'ancienne Bibliothèque nationale n'avait pas attendu l'EPBF pour se lancer à la conquête de nouveaux moyens de diffusion du patrimoine : on peut citer pour exemple la réalisation du vidéodisque *Images de la Révolution Française* à l'occasion du Bicentenaire⁶, sans doute l'une des premières grandes entreprises englobant numérisation et valorisation d'un corpus d'images, dont l'idée émerge dès 1982. Enfin, la date de 1992, qui voit le Code du Patrimoine ajouter à la collecte par dépôt légal les documents électroniques, logiciels et bases de données, mérite également d'être notée comme un tournant dans le périmètre des collections.

Cette histoire du premier âge numérique de la BnF, qui commence en 1982, en 1988 ou en 1992, mériterait d'être racontée. Rarement traitée en profondeur dans les nombreuses publications qui détaillent la mutation organisationnelle que connaît la BnF dans la dernière décennie du vingtième siècle, elle se focaliserait sans doute, bien plus que nous ne pouvons le faire ici, sur la naissance des collections audiovisuelles et multimédias. Il faudrait alors remonter jusqu'aux années 1970, voire au-delà⁷, pour les embrasser dans leur ensemble, ce qui dépasse très largement l'ambition qu'il nous était possible de nous donner dans le cadre de notre doctorat sur travaux.

À la recherche d'un élément disruptif plus proche et peut-être plus significatif au regard de la nature des transformations qui ont affecté l'établissement depuis l'an 2000 environ, nous avons donc porté notre regard sur la naissance du web. En effet, ce « numérique » qui affecte aujourd'hui les documents que nous produisons et consultons, nos processus de travail, de communication, nos démarches administratives et d'une façon générale, tous les aspects de notre vie, devient un média grand public au milieu des années 1990. En 1996, La BnF pro-

6 Bruckmann, Denis. « Le Vidéodisque images de la Révolution française ». *Bulletin des bibliothèques de France*, 1990, n° 2, p. 122-135. En ligne : <http://bbf.enssib.fr/consulter/bbf-1990-02-0122-006>.

7 Le dépôt légal est élargi à l'édition phonographique en 1938 et en 1977 aux vidéogrammes et documents multimédias. Cf. <https://www.bnf.fr/fr/departement-de-laudiovisuel#bnf-historique-du-d-partement>.

pose un premier site web et la fondation Internet Archive, aux États-Unis, commence à se préoccuper de la constitution d'une mémoire des contenus qui sont publiés en ligne. En 1997, le projet de numérisation de la BnF trouve pour la première fois une concrétisation sur le réseau, avec la toute première version de Gallica, sa bibliothèque numérique.

C'est donc la date de 1997 que nous avons retenue, car ce jalon – bien qu'imparfait – semble marquer le début de l'ère au cours de laquelle le web aura un impact décisif sur l'établissement. Cet impact ne cessera de se préciser, se renforcer et se modifier, au fur et à mesure que la BnF acquerra l'expertise et la maîtrise nécessaires pour devenir un véritable acteur de ce nouvel environnement.

Cette extension chronologique, assortie de la préoccupation d'éviter les biais liés à une implication trop personnelle dans les événements relatés, nous a conduite à mobiliser différents outils méthodologiques pour aborder la question de la naissance à la BnF d'un nouveau patrimoine numérique.

Une approche inspirée des sciences de l'information et des sciences que l'on avait coutume d'appeler « auxiliaires » de l'histoire (paléographie, codicologie, diplomatique, etc.) conduit à travailler sur la « matérialité » du numérique comme on a pu travailler auparavant sur celle du livre⁸. Dans une première partie qui coïncide avec une période d'exploration et de découverte pour la BnF entre 1997 et 2008, nous tenterons ainsi de border la notion de document numérique et son rapport à la bibliothèque.

Des éléments tirés de l'anthropologie et de la sociologie sont également utiles pour comprendre comment l'humain est affecté par la transformation numérique. Dans notre deuxième partie, qui porte sur l'affirmation de la notion de « collections numériques » entre 2008 et 2018, nous combinons ainsi considérations sur l'organisation de la bibliothèque et son évolution d'une part, et observations sur les nouveaux usages induits notamment dans le champ de la recherche.

Enfin une approche historique reste indispensable pour comprendre comment les évolutions de la période ont suscité prises de consciences, nouvelles stratégies et évolutions organisationnelles, jusqu'à faire émerger l'idée de patrimoine numérique. Les sources que nous avons réunies pour appréhender ces faits ont été principalement des productions documen-

8 Henri-Jean Martin (dir.), *La naissance du livre moderne : XIV^e-XVII^e siècles : mise en page et mise en texte du livre français*. Paris : Éditions du Cercle de la Librairie, 2000.

taires et publications de l'époque, aussi bien les nôtres dans l'esprit du doctorat sur travaux que celles d'autres professionnels de la bibliothèque, mais aussi d'autres institutions et au-delà. En effet, si le centrage de notre propos sur le cas particulier de la BnF est inévitable du fait de notre positionnement et de la nature de l'exercice, il s'inscrit dans un contexte plus vaste, dans un mouvement d'ampleur internationale et globale, que nous ne ferons qu'esquisser. Nous nous sommes donc efforcée de mettre l'expérience de la BnF en perspective, même si notre objet principal est bien d'étudier son histoire.

L'idée de processus de patrimonialisation, ses caractéristiques, ses modes de fonctionnement, ses étapes, a été largement étudiée, en premier lieu s'agissant des monuments et des collections muséales, et plus récemment dans de nombreux autres champs, incluant le patrimoine immatériel et le patrimoine écrit. Elle implique la considération d'un temps long, qui remonte à la Révolution française pour l'émergence de la notion de patrimoine, connaît un bouleversement considérable avec les *Lieux de mémoire* de Pierre Nora dans les années 1980⁹, et continue de se développer aujourd'hui. Ce champ de recherche se situe à l'articulation de plusieurs disciplines : histoire, histoire de l'art, anthropologie, sociologie, sciences de l'information... qui apportent autant d'approches potentielles pour le questionnement que nous avons souhaité soulever dans notre troisième partie : le processus de patrimonialisation s'applique-t-il au numérique de la même façon qu'à d'autres objets matériels ou immatériels ? C'est à travers une courte histoire, sur vingt ans, des collections de la BnF, confrontée au prisme de notre propre expérience et mémoire, que nous proposons d'ouvrir ce débat. Nous ne doutons pas qu'il ne s'agit là que des premières pages d'un récit encore en train de s'écrire.

9 Paris : Gallimard, 1984, 3 volumes.

Première partie :

Technique et matérialité du document numérique,
l'invention d'un outillage nouveau (1997-2008)

Les documents numériques présentent la particularité de nécessiter la médiation d'un matériel, d'un outil pour accéder au contenu. Avant eux, les documents audiovisuels – 78 tours, cassettes vidéo, etc. – exigeaient déjà un matériel de lecture spécifique, sans lequel on ne pouvait accéder à l'information inscrite sur le support. Avec le CD-rom ou les jeux vidéo entrés dans les collections de la BnF à partir de 1992¹⁰, ces systèmes appartenaient déjà au monde numérique. Pour ces documents d'un genre particulier, la relation entre le lecteur et le document passe par un artefact, une machine dont les caractéristiques techniques (comme la qualité de restitution) et ergonomiques (les moyens d'interaction entre l'homme et la machine) sont susceptibles d'affecter l'expérience d'accès au contenu. Cette problématique n'est donc pas totalement nouvelle lorsque la question de la création d'une bibliothèque numérique se pose dès les années 1990 : cette bibliothèque « d'un type entièrement nouveau », imaginée par Jacques Attali et annoncée par François Mitterrand le 14 juillet 1988¹¹. Pourtant, la numérisation pose une question inédite : les logiciels et matériels permettant d'accéder au contenu sont désormais à la main de la bibliothèque. Il lui revient d'imaginer dans quel environnement, sous quelle forme et avec quelles fonctionnalités elle va construire la relation entre usager et documents. Dans les premières années de conception de l'EPBF, cet enjeu s'exprime à travers le projet de PLAO (Poste de Lecture Assisté par Ordinateur), qui doit répondre à la nature particulière du document numérique en proposant, pour les salles de la future grande bibliothèque, une station de lecture dotée de capacités pointues de recherche et de manipulation des textes numérisés.

Cependant, l'irruption du web va transformer le défi du numérique tel qu'il avait été pensé à l'origine : d'une part, en soulevant la question de la mise en ligne des documents numérisés par la bibliothèque, et d'autre part, en faisant apparaître d'innombrables ressources documentaires accessibles sur le réseau, face auxquelles elle doit trouver comment jouer son rôle. En 1994, au moment de la fusion entre l'établissement public préfigurateur de la nouvelle bibliothèque et l'ancienne Bibliothèque nationale, le PLAO est abandonné, au profit

10 Avec la loi n° 92-546 du 20 juin 1992 relative au dépôt légal, qui l'étend aux documents multimédias, logiciels et bases de données. En ligne : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006079299>.

11 Ce terme n'apparaît en fait pas dans le discours du 14 juillet, mais dans la lettre de mission adressée au premier ministre quelques semaines plus tard. Cf. Jean Gattegno, *La Bibliothèque de France à mi-parcours : de la TGB à la BN bis ?* Paris : Éditions du Cercle de la librairie, 1992, p.16-17.

d'un nouveau projet de numérisation qui veut toucher un public plus large, sur place et à distance¹².

Dès lors, à la fin des années 1990 et au début des années 2000, deux questions se posent pour la bibliothèque : d'une part, celle de son positionnement s'agissant de l'utilisation du web comme canal de diffusion de ses collections, d'autre part, celle de l'inclusion de documents numériques créés par d'autres et diffusés en ligne dans ces mêmes collections, notamment au titre du dépôt légal. Nous proposons d'étudier dans cette première partie comment, au fil d'une décennie, la BnF a acquis la maîtrise du web en tant qu'espace documentaire, mobilisant pour ce faire de nouveaux outils.

12 Gaëlle Béquet, *Trois bibliothèques européennes face à Google*. Paris : École nationale des chartes, 2014, p.111.

I. S'appropriier les codes du web (1997-2003)

En 1997, une première version de Gallica voit le jour avec 2.500 imprimés libres de droit et 10.000 images. Malgré sa construction artisanale, l'accueil enthousiaste dont elle fait l'objet encourage le président de la BnF, Jean-Pierre Angremy, à annoncer dès l'année suivante une nouvelle version nommée « Gallica 2000 », dotée d'une politique scientifique plus claire, et accompagnée d'une démarche de valorisation. Celle-ci prendra la forme de dossiers thématiques : Gallica classique, puis Voyage en France¹³.

Au moment même où s'élaborait la première version de la bibliothèque numérique française, la fondation américaine Internet Archive entreprenait de créer une mémoire de la Toile mondiale en archivant les pages web par collecte automatique, à l'aide de robots¹⁴. La transition d'un certain nombre de publications scientifiques vers un mode en ligne poussait déjà les bibliothèques nationales à se poser la question de la continuité du dépôt légal, qui s'était adapté par le passé à différents types de médias et de publications. Pourtant, l'émergence d'un besoin de conservation de ce réseau par nature volatile n'était en rien une évidence : d'autres réseaux avaient précédé le web, à commencer en France par le Minitel, sans pour autant susciter une telle prise de conscience ni une action en ce sens. Dans la deuxième moitié de la décennie 1990, l'avenir florissant du web ne fait pas l'unanimité : beaucoup doutent de sa capacité à être adopté par le grand public, de la possibilité d'y développer des modèles économiques ou encore de sa fiabilité technique, allant jusqu'à penser qu'il s'agit d'une mode passagère¹⁵.

Il s'agit donc d'un pari lorsqu'en 1999, la BnF crée un « département de la bibliothèque numérique », chargé de toutes les questions relatives au web, distinct du département des systèmes d'information, en charge d'une informatique qui apparaît comme plus stable et pérenne. De Gallica au site institutionnel en passant par le projet d'archivage de l'Internet, ce département est celui où se joueront les premières expérimentations sur le numérique. L'enjeu est immense : il s'agit de comprendre et s'appropriier un environnement nouveau, qui boule-

13 G. Béquet, *op. cit.* p.137.

14 La première collecte est réalisée en 1996. Cf. Francesca Musiani, Camille Paloque-Bergès, Valérie Schafer, *et al. Qu'est-ce qu'une archive du web ?* Marseille : OpenEdition Press, 2019, p.10.

15 Valérie Schafer, *En construction : la fabrique française d'Internet et du Web dans les années 1990*. Bry-sur-Marne : INA, 2018, p. 26.

verse la relation de la BnF à ses collections, à son public, à son outil de travail, et même à ses missions : sans doute la « plus grande révolution depuis Gutenberg »¹⁶.

A. Ubiquité : la bibliothèque face aux usages du web

Pour la bibliothèque, l'irruption du web signe la fin d'une maîtrise de l'espace à la fois physique et technique dans lequel le document était accessible. Là où le PLAO était conçu comme une station de travail complète, réunissant dans un environnement clos les outils nécessaires au chercheur expert capable d'effectuer des opérations avancées sur les textes dématérialisés¹⁷, en 1997 Gallica est une bibliothèque numérique ouverte sur le web.

Pour la BnF, le changement est de taille. Le public qui fréquente les salles de lecture est connu, identifié et formé. Il est également « captif », au sens où il doit, pour accéder aux contenus dont il a besoin, se plier aux règles fixées par la bibliothèque, seule détentrice du savoir-faire documentaire. Sur le web, au contraire, la bibliothèque se trouve confrontée à un public très large et globalement peu identifié, qui lui impose des impératifs implicites propres aux usages du web : attractivité graphique, temps de réponse et de chargement rapides, ergonomie globale... Il devient impossible de compter sur le savoir-faire documentaire de l'internaute, de le former ou même de l'informer, la lecture d'un mode d'emploi sur un site web étant loin d'être une évidence. Les codes et règles de ce nouvel espace virtuel échappent totalement aux méthodes héritées de la tradition documentaire, comme à celles qui relèvent de la démarche scientifique classique¹⁸.

Pour s'adapter à ce nouveau contexte, la bibliothèque doit faire un effort considérable à la fois sur le plan technique (mettre en place des serveurs performants, développer une architecture web) et sur le plan des usages (comprendre les standards de fait d'Internet et les adopter). Cet effort prend le pas, pendant un temps, sur celui qui consistait à imaginer les nouveaux modes de recherche susceptibles de découler de la nature numérique des documents. En

16 Depuis les années 2000, on trouve cette expression sous la plume de nombreux auteurs, par exemple Jean-Noël Jeanneney dans son *Histoire des médias* (Paris : Seuil, 2011).

17 G. Béquet, *op.cit.*, p.72-73.

18 Cf. **Emmanuelle Bermès, « Disintermediation and re-intermediation: the challenge of digital data for heritage professions » dans B. Dufrene (dir.), *Heritage and digital humanities : how should training practices evolve ?* Zurich, Berlin : Lit, 2017. Textes issus du colloque Patrimoines et humanités numériques organisé par le Labex Arts H2H (Universités de Paris 8 et Paris Ouest) en collaboration avec les Archives nationales et la Bibliothèque nationale de France.

1997 d'abord, puis avec la deuxième génération appelée « Gallica 2000 », la bibliothèque numérique se concentre donc sur des fonctions d'accès de base. La typologie documentaire est réduite à l'essentiel, dans l'héritage de la période précédente qui catégorisait le numérique en deux grands ensembles, le texte d'un côté, l'image de l'autre.

« *Comprendre le web des années 1990 implique de se défaire de notre vision faite d'hyperconnectivité, de forfaits illimités, de navigation fluide, de moteurs de recherche puissants, qui créent un environnement numérique bien différent de celui que connaissent les premiers internautes* » nous rappelle Valérie Schafer¹⁹. La proposition de la BnF, avec Gallica 2000, doit tenir compte des conditions – loin d'être optimales – qui sont celles de l'environnement des internautes et tenter de s'y adapter : pour cela, mieux connaître leurs besoins est essentiel. De 2001 à 2003, la BnF s'engage dans un partenariat avec France Télécom pour étudier les usages de Gallica, dans le contexte plus large du web et d'Internet : le projet Bibusages²⁰. Une centaine de volontaires sont équipés de systèmes de traçage qui enregistrent la totalité de leur navigation, que celle-ci concerne Gallica ou non. L'enquête positionne ainsi Gallica en regard des moteurs de recherche, des sites de « contenus » comme les sites de presse, ou encore des sites de commerce en ligne de biens culturels. Elle s'efforce de qualifier les pratiques documentaires dans cet espace global d'information caractérisé par son absence de frontières thématiques, géographiques ou temporelles.

Par sa démarche même, qui ne se limite pas à l'étude de Gallica mais englobe l'ensemble de la navigation hypertextuelle des internautes du panel, l'étude Bibusages marque une prise de conscience : la bibliothèque n'est pas seule sur le web, les contenus qu'elle propose sont en concurrence avec d'autres, émanant de la sphère privée. C'est la fin d'un univers de la rareté²¹ dans lequel les bibliothèques étaient, pour toute recherche, un passage obligé, parce qu'elles étaient les seules à collecter et donner accès de manière ouverte, globale et gratuite aux ressources documentaires de toute nature. La coexistence de l'offre documentaire de

¹⁹ *En construction... op.cit.*, p. 31.

²⁰ T. Cloarec, « BibUsages : une étude sur les pratiques des utilisateurs de Gallica » dans *Culture et recherche* n°100, janvier-février-mars 2004. En ligne : <http://www.culture.gouv.fr/var/culture/storage/lettre-recherche/cr100.pdf>.

²¹ Expression utilisée par Bertrand Calenge dans *Les bibliothèques et la médiation des connaissances*, Paris : Éditions du Cercle de la librairie, 2015, p. 11.

la BnF avec d'autres sources de contenus participe de l'appréhension du web comme un espace documentaire, dans lequel la bibliothèque doit trouver sa place.

B. Volatilité : la bibliothèque face à un médium impermanent

À une époque où les ordinateurs ne sont pas connectés en permanence au réseau et où l'on paye son abonnement à l'heure, le projet Bibusages fait apparaître les limites techniques de l'idée de bibliothèque numérique : la lecture en ligne est pratiquement inexistante, le temps de navigation est essentiellement consacré à la recherche d'informations. Pour la majorité des usagers, l'une des pratiques dominantes est alors de télécharger les documents numérisés en intégralité pour recréer sur leur ordinateur personnel leur propre bibliothèque. Mais la volatilité des contenus sur le web n'est pas seulement liée à des difficultés techniques d'accès : elle est intrinsèque à ce médium qui se comporte comme un palimpseste, que les fournisseurs de contenus ne cessent de faire évoluer, où les liens se brisent, où les textes que l'on pouvait lire sur la page d'un journal en ligne deviennent payants dans les archives le lendemain. La bibliothèque est elle-même confrontée à cette impermanence lorsque les acquisitions de revues scientifiques, autrefois pérennes à l'envi puisqu'elle était libre de désherber ou non les volumes papier, deviennent des « locations », des acquisitions temporaires de licences d'accès, susceptibles d'être révoquées par l'éditeur si les budgets d'acquisitions viennent à manquer.

À la BnF, le département de la bibliothèque numérique emmené par Catherine Lupovici et avec elle, Julien Masanès alors en charge du projet d'archivage de l'Internet, se pose très tôt la question de la conservation de ces publications qui se déplacent progressivement vers le web. D'un point de vue technique, la faisabilité du projet est loin d'être une évidence, mais plus encore, la légitimité politique, juridique et patrimoniale de l'idée même d'archivage du web reste encore à démontrer. Deux leviers sont alors utilisés pour convaincre de la nécessité d'une action dans ce domaine : d'une part la continuité des collections, en faisant des parallèles entre ce que la bibliothèque conservait dans le monde du papier et ce qu'elle collecte dans celui du web, d'autre part la dimension éphémère des sites, qui justifie leur archivage

d'autant plus qu'elle est associée à un caractère d'intérêt scientifique, historique et sociologique évident²².

Ainsi la première collecte expérimentale du web réalisée par la BnF en 2002 porte sur les sites électoraux : parce que la bibliothèque dispose d'un historique de collecte d'éphémères politiques, parce que la volatilité des sites de campagne est facile à démontrer, enfin parce que la qualité politique et potentiellement historique du sujet facilite la projection de l'utilisation de ce matériau comme source par de futurs chercheurs, donc la perspective d'un usage. Cette collecte électorale, répétée à chaque échéance politique depuis lors, sera l'une des premières à faire l'objet d'une étude par des élèves de Sciences Po, encadrés par Dominique Cardon, en partenariat avec la BnF²³.

C'est également cette question de l'usage par les chercheurs, travaillée dans un esprit de continuité avec la collection patrimoniale traditionnelle, qui contribue en 2003 à faire évoluer la bibliothèque numérique dans le sens de plus de stabilité. Le projet IDA²⁴ lancé en 2003 visait la numérisation des manuscrits autographes de Proust et Flaubert, que les chercheurs de l'Item, laboratoire de l'ENS attaché notamment à l'analyse génétique des textes littéraires, ambitionnaient de doter d'un dispositif de transcription augmentée inspiré du projet franco-allemand *HyperNietzsche*²⁵. On retrouve ici l'enjeu de rendre possibles des usages de recherche avancés, qui sous-tendait le projet de PLAIO, mais cette fois, sur le web, ce qui implique d'une part la numérisation de documents complexes, et d'autre part l'utilisation d'identifiants pérennes garantissant la citabilité des documents numérisés et mis en ligne.

La numérisation des manuscrits contemporains fait émerger des questions nouvelles : des brouillons de Flaubert rédigés tête-bêche aux paperoles de Proust qui se déploient parfois sur plusieurs dizaines de centimètres, l'enjeu est la transposition à l'écran de structures com-

22 Voir l'introduction de Magali Haettiger à son article de 2003 : « Vers la conservation des sites web régionaux ». *Bulletin des bibliothèques de France*, 2003, n° 4, p. 77-84. En ligne : <http://bbf.enssib.fr/consulter/bbf-2003-04-0077-015>.

23 *Séminaire IEP 2006 : description et organisation du séminaire de Dominique Cardon et Gildas Illien. Internet en campagne : les usages des archives électorales de l'internet de la BnF*. Archives numériques de la BnF : BnF-ADM-2011-073722.

24 Pour « Interchange of Data between Administrations » du nom de l'appel à projet. Voir <https://c.bnf.fr/DXV>. Le projet a trouvé une suite en 2006-2008 avec un autre projet nommé Optima : <https://c.bnf.fr/DXY>.

25 Paolo d'Iorio, *HyperNietzsche. Modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines. Questions philosophiques, problèmes juridiques, outils informatiques*. Paris : PUF, 2000. En ligne : <http://www.diorio.info/puf/>.

plexes que le chercheur doit retrouver à l'identique de l'original. Le travail des généticiens sur les manuscrits littéraires, notamment la transcription, implique la définition de zones précises à l'intérieur d'une image numérisée. La pérennité des liens sur le web à un niveau aussi fin modifie le rapport entre le document original et le document numérisé. Il devient impossible de remplacer un fichier numérisé par un autre, même si le fichier d'origine était défectueux, ou de substituer la numérisation d'un exemplaire à un autre.

Rappelons qu'à cette époque, l'essentiel des ouvrages et revues envoyés en numérisation n'étaient pas issus des collections de la bibliothèque, mais étaient rachetés à cette fin sur le marché de l'antiquariat²⁶ : seul importait l'accès au contenu. Les projets de numérisation de documents spécialisés comme les manuscrits contemporains font donc réapparaître la préoccupation de la matérialité du document que l'on numérise. De cette question abordée sous l'angle de la citabilité, c'est-à-dire des usages de recherche appliqués à un objet dont on entérine son mode d'accès en ligne, découle une conception structurante de la bibliothèque numérique. On s'éloigne d'une conception du livre comme « matériau contenant des informations »²⁷ qui prévalait depuis 1990. Avec la numérisation des collections spécialisées, le document numérisé devient, en lui-même, un objet d'étude : en somme, un nouvel original, doté de ses propres caractéristiques matérielles²⁸.

C. Déstructuration : la bibliothèque entre reproduction et réédition

À l'exemple de la radio ou de la télévision, le document numérique fait l'objet d'une transmission qui le dissocie à la fois de la notion de support – l'utilisateur a entre les mains le terminal qui permet d'accéder au document, et non plus le document lui-même – et de la tem-

26 Un test réalisé en 1990 sur 600 ouvrages pour comparer la numérisation des originaux à celle des ouvrages massicotés avait montré que les risques de détérioration étaient importants « pour certains types de reliure » et que le coût était bien moindre pour numériser des livres massicotés, puisqu'il n'y avait pas besoin d'opérateur. La décision fut donc prise de se limiter à la numérisation des microformes et d'ouvrages acquis à cette fin pour être massicotés, décision qui ne fut pas remise en cause avant le lancement de la numérisation de masse en 2006-2007. Cf. Thierry Boucher, *La numérisation des imprimés à la Bibliothèque nationale de France*. Mémoire d'études pour le diplôme de conservateur de bibliothèques, ENSSIB, 1994, p. 16-17. En ligne : <https://www.enssib.fr/bibliotheque-numerique/documents/62891-la-numerisation-des-documents-imprimes-a-la-bibliotheque-nationale-de-francememoire-d-etude.pdf>

27 F. Stasse, *op. cit.*, p. 56.

28 Sur cette question, voir Mélanie Roustan (dir.), *La recherche dans les institutions patrimoniales : sources matérielles et ressources numériques*. Villeurbanne : Presses de l'Enssib, 2016.

poralité de la lecture sous forme de flux – l'une des caractéristiques de l'hypertexte étant la délinéarisation des objets auxquels il permet d'accéder²⁹. Si le web change véritablement la donne, c'est parce qu'il remet en cause la structure même du document, la relation entre le tout et ses parties, et le prive de ses limites, d'une « finitude » qui était apportée par le support physique et vole en éclats avec la navigation de lien en lien sur le réseau.

Cette nature particulière du web représente un défi pour la création d'une archive des contenus qui y circulent. Si la plupart des ressources présentes sur le web peuvent être assimilées ou affiliées à des documents que la bibliothèque a collectés et conservés sur support auparavant, la nature décentralisée et fragmentaire du web se révèle notamment lorsqu'on tente de quantifier le matériau à collecter³⁰. Les sites et pages sont constitués chacun de dizaines de fichiers indépendants ; la logique de l'hypertexte permet de combiner ces éléments sans qu'il soit possible de prédire à quel moment on quitte un site donné. Face à ce « puzzle » qu'est le web, il s'avère très difficile de délimiter les bornes de quelque chose que l'on pourrait appeler un « document ».

L'impossibilité de faire correspondre l'entité intellectuelle, ou structure logique, à la réalité matérielle, ou structure physique, est une dimension essentielle du numérique et du web qui oblige la bibliothèque à adopter des positionnements radicaux pour traiter de tels objets. Dès 2002, la première collecte du web électoral est automatisée³¹. Des « graines » ou listes d'URL sont fournies en entrée au robot pour la collecte, puis celui-ci parcourt les pages web de lien en lien. Les différents fichiers qui constituent chaque page sont enregistrés sans ordre particulier dans un format conteneur appelé (W)ARC³², jusqu'à atteindre une volumétrie donnée. Les gigaoctets ou téraoctets qui en résultent sont aux archives du web ce que les mètres et kilomètres linéaires sont aux magasins de la bibliothèque ; un magasin dans lequel il serait impossible de raccrocher les unités physiques à des objets documentaires intellectuels comme

29 Jean-Michel Salaün, *Vu, lu, su : les architectes de l'information face à l'oligopole du web*. Paris : La Découverte, 2012, p. 69 et ss.

30 Cf. **Emmanuelle Bermès et Gildas Illien, « Metrics and Strategies for Web Heritage Management and Preservation ». IFLA WLIC 2009 : IFLA World Library and Information Congress, 75th IFLA General Conference and Assembly. Milan, 2009. En ligne : <https://www.ifla.org/past-wlic/2009/92-bermes-en.pdf>

31 On utilise à l'époque le robot ou « crawler » HTTrack.

32 Le format ARC – abréviation d' « archive » – créé par Internet Archive a été utilisé à la BnF pour la collecte des archives web jusqu'en 2014, avant d'être remplacé par le format WARC, similaire mais plus performant et normalisé à l'ISO en mai 2009 sous le numéro ISO-28500.

le « site » ou la « page » web. Dès lors, il devient difficile de recourir à des techniques de traitement de collection classiques comme, au premier chef, le catalogage, dont la logique même est de faire correspondre la description d'un objet matériel avec celle de son contenu.

D'un point de vue herméneutique, les chercheurs qui se sont penchés par la suite sur les archives de l'Internet ont posé de manière claire l'idée que « *l'archive d'un site n'est pas le site archivé* »³³, ni même une « trace » au sens où l'archive traditionnelle est la trace de l'activité qui l'a produite. Le matériau mis à leur disposition par les bibliothécaires est une « reproduction », une « copie » du web qu'il convient d'approcher avec un regard critique. En ce sens, l'archive web peut être rapprochée de la numérisation, qui soulève le même type de problèmes, notamment s'agissant du conflit entre structure physique et structure logique.

Le projet de numérisation originel et son corollaire en termes d'accès, le PLAO, visaient ainsi en premier lieu la mise à disposition de texte : c'est le contenu des documents, donc leur structure logique, qui faisait l'objet d'un acte de reproduction. Cependant, il était impossible d'atteindre à un coût abordable les volumétries envisagées³⁴ ; de plus, produire un flux de texte linéaire devenait rapidement insuffisant dès qu'on avait besoin de convertir au numérique des documents complexes, par exemple des dictionnaires, des encyclopédies, des revues structurées en colonnes, tableaux ou articles, des ouvrages contenant des illustrations scientifiques, etc. Face à un tel constat, la position de la BnF a été dans un premier temps de privilégier la reproduction en mode image de la structure physique des documents, c'est-à-dire les volumes et les pages, à l'exclusion d'autres formes de hiérarchie textuelle : par exemple, les articles d'une revue ont une structure logique qui ne correspond pas nécessairement à cette structure physique (ils peuvent commencer au milieu d'une page, s'étendre sur plusieurs pages...)³⁵. Dans un article qui met en avant l'importance du mode texte pour l'accessibilité,

33 Bruno Bachimont, « L'archive du web : une nouvelle herméneutique de la trace. » Dans *Web Corpora*, 21/06/2017. En ligne : <https://webcorpora.hypotheses.org/288>

34 Les autres projets qui avaient choisi cette voie, comme le projet Gutenberg, s'appuyaient sur la force de travail d'une communauté en ligne, ou plus modestement, comme dans le cas de la bibliothèque numérique de Lisieux, sur un seul contributeur : <http://www.bmlisieux.com/>. Cf. « Les choix techniques à Lisieux : des réserves à l'Internet, les étapes de la mise en ligne », dans Charlette Burési et Laure Cédelle-Joubert (dir.), *Conduire un projet de numérisation*. Villeurbanne : Presses de l'Enssib, 2002, coll. « Boîte à outils », n° 13.

35 D'autres initiatives comme le portail Persée, initié en 2003 par le Ministère de l'enseignement supérieur et de la recherche, adoptèrent au contraire cette approche logique, à la faveur d'une collection beaucoup plus homogène. Cf. Nathalie Fargier, « Persée, une bibliothèque numérique par et pour les chercheurs. » Dans *La francesistica italiana à l'ère du numérique*, Université de Gênes, SUSLLF (Società Universitaria per gli

Jean-Didier Wagneur³⁶ expose les motivations qui ont sous-tendu ces choix et évoque les tentatives effectuées pour les dépasser :

Nous avons le devoir patrimonial de restituer l'image du document tel qu'il a été déposé et le choix du mode image (fac-similé électronique) s'est imposé. (...) Depuis cette première étape, et toujours dans l'esprit d'une bibliothèque numérique, nous avons développé d'autres programmes en mode texte en accord avec notre approche documentaire. (...) Il est très vite apparu nécessaire de s'orienter vers la saisie de dictionnaires très sollicités par les chercheurs, mais échappant au secteur de l'édition pour des raisons économiques. (...) Ces dictionnaires (...) font partie des outils spécifiques des salles d'orientation bibliographique des bibliothèques.

Les expérimentations évoquées portaient sur la numérisation intégrale en TEI³⁷ de deux dictionnaires biographiques, le *Dictionnaire universel des contemporains* de Gustave Vapereau et le *Dictionnaire historique* de Moreri, et d'une revue, la *Revue de synthèse*. Un projet d'édition scientifique ambitieux fut également entrepris avec l'ITEM³⁸. Prenant pour point de départ une transcription en TEI du manuscrit du roman *Le rêve* de Zola, il fut mis en ligne en octobre 2002³⁹. Ces incursions dans une approche logique du texte restèrent cependant isolées, parce qu'elles se heurtaient à l'ambition de numérisation en nombre qui sous-tendait toujours Gallica. Le principe d'un encodage TEI fut par la suite généralisé aux seules tables des matières, ce qui permettait de structurer un plus grand nombre d'ouvrages, de manière plus systématique. La BnF se rangea donc *de facto* à l'analyse rendue déjà plus de dix ans auparavant par les fondateurs de la TEI⁴⁰ : la complexité de l'objet apparemment simple

Studi di Lingua e Letteratura Francese), Institut français d'Italie, 2014, Gênes, Italie. En ligne : <https://hal.archives-ouvertes.fr/hal-01187830/>

36 Jean-Didier Wagneur, « Gallica : la bibliothèque électronique de la BnF : quel accès pour les personnes handicapées visuelles ? » dans *Bibliothèques et publics handicapés visuels*. Paris : Éditions de la Bibliothèque publique d'information, 2002. En ligne : <http://books.openedition.org/bibpompidou/1500>.

37 Text Encoding Initiative : la TEI est l'un des principaux systèmes d'encodage de texte basés sur les langages à balise utilisés en SHS depuis 1987.

38 Institut des Textes et Manuscrits Modernes.

39 Jean-sébastien Macke, « Une édition en ligne : *Le Rêve* de Zola sur Gallica. » Dans *Genesis (Manuscrits-Recherche-Invention)*, numéro 23, 2004. p. 169-170. En ligne : https://www.persee.fr/doc/item_1167-5101_2004_num_23_1_1307_t1_0169_0000_2.

40 Steven DeRose, David Durand, Elli Mylonas, Allen Renear, « What is text, really ». Dans *Journal of Computing in Higher Education*, déc. 1990, Volume 1, tome 2, p. 3-26 ; et David Durand, Elli Mylonas et Allen Renear, « Refining our notion of what text really is » (1993, en ligne :

qu'est le texte et la subjectivité de l'acte de transposition de ce contenu au numérique. En renonçant à l'acte d'interprétation qui découle de la représentation de la structure logique, elle se range pour de bon du côté de la reproduction.

Qu'elle collecte le web, qu'elle numérise en mode texte ou en mode image, la bibliothèque se trouve ainsi dans les premières années du siècle confrontée à une interrogation fondamentale : en exerçant ces nouvelles activités, agit-elle dans la continuité de ses missions de conservation et de mise à disposition du patrimoine, ou crée-t-elle de nouveaux objets ? Est-elle toujours bien dans son rôle de bibliothèque ? La notion de « copie », de « reproduction », de « transposition au numérique » voire de « dématérialisation » ne manque pas de soulever des problèmes similaires à ceux que l'audiovisuel avait déjà eu à résoudre en migrant les contenus d'un support à l'autre pour des raisons de conservation. La différence majeure réside dans le fait que cette migration de support, pour l'audiovisuel, a été immédiatement envisagée comme un plan de sauvegarde, en raison de l'obsolescence rapide des supports d'origine, ce qui n'est pas le cas de la numérisation des imprimés⁴¹.

Avec le document numérique, la bibliothèque crée un ensemble d'objets ambigus du point de vue de leur usage. La numérisation vise à l'origine à amplifier la dissémination de la connaissance, mais la reproduction des documents spécialisés pose la question de son rôle de conservation – par la mise à disposition d'un support de substitution, à des fins de préservation de l'original. Avec le PLAO puis les projets de numérisation en TEI, la bibliothèque vise à permettre le développement de nouveaux usages autour des textes, mais la réalité des coûts et de la nature des collections l'engage sur la voie de la création de simples fac-similés en mode image. Pour créer une mémoire du web, elle tente de l'archiver tout en admettant que

<http://cds.library.brown.edu/resources/stg/monographs/ohco.html>) : la notion de texte est définie comme « une hiérarchie ordonnée d'objets de contenus », mais le deuxième article pose la question de la superposition des hiérarchies différentes au sein d'un même texte et admet que la perspective de l'encodeur a une influence déterminante sur le résultat obtenu.

41 Dans l'ouvrage de Daniel Renoult et Jacqueline Sanson en 2001, la numérisation est présentée comme une stratégie de conservation pour les documents audiovisuels, alors que pour le livre il n'est question que de microfilmage. *La Bibliothèque nationale de France : collections, services, publics*. Paris : Éditions du Cercle de la librairie, 2001, p. 143 et 145.

« *l'idéal d'une archive isomorphe est impossible à atteindre* »⁴². Paradoxe supplémentaire : pour des raisons de droits, cette archive de l'Internet n'est accessible que dans les emprises de la bibliothèque, alors qu'elle reproduit un média numérique d'accès libre et gratuit en tout lieu.

Ces expérimentations, ces tâtonnements, montrent qu'à cette époque la perception de ce que sont les documents numériques est encore en construction. La bibliothèque, en tant qu'institution patrimoniale, doit à la fois trouver son rôle face à ces nouveaux objets et définir des processus technologiques gérables pour l'exercer. Or, la nature du web et des objets que la bibliothèque y manipule crée une dissension entre ses savoir-faire traditionnels et l'impossibilité de les appliquer tels quels dans ce contexte.

Les premières années du XXI^e siècle sont ainsi marquées par la prise de conscience de l'ampleur du changement induit par le numérique et le web dans les pratiques documentaires : un changement dont on pressent alors qu'il n'en est qu'à ses balbutiements. En 2002, François Stasse évoque l'enjeu de « *quitter l'ère Gutenberg* » : comme au temps des incunables, les premières tentatives d'exploitation de la technologie nouvelle ont cherché à se raccrocher à des objets connus, à reproduire des formes identifiées avant de se stabiliser pour pouvoir évoluer. En 2003, le CNRS lance plusieurs « réseaux thématiques de recherche » (RTP) dont l'un d'eux, animé par Jean-Michel Salaün, a pour mission de se pencher sur la notion de document numérique. Le RTP-doc se dote d'un pseudonyme, Roger T. Pédauque, pour écrire collectivement un document fondateur et interdisciplinaire sur la question. Celui-ci réaffirme l'idée que l'on fait face à une évolution dont l'ampleur est comparable à celle de l'invention de l'imprimerie et qui doit faire l'objet d'une réflexion transdisciplinaire. Il propose une définition du document numérique suivant trois axes, selon que l'on envisage le document comme un objet matériel, comme un contenu porteur de sens ou dans ses fonctions sociales de trace ou de preuve, notamment à travers l'archivage⁴³. Trois axes qui font écho aux problématiques

42 Julien Masanès, « L'archivage du web ». Dans Lisette Calderan, Bernard Hidoine et Jacques Millet (dir.), *Pérenniser le document numérique*. Séminaire INRIA, 2-6 octobre 2006, Amboise. Paris : ADBS, 2006, p. 101-122.

43 Le premier des trois articles qui composent ce travail fut publié dans HAL en 2003 : Roger T. Pédauque. *Document : forme, signe et médium, les re-formulations du numérique*. 2003. En ligne : https://archivesic.ccsd.cnrs.fr/sic_00000511/. L'ensemble des productions du RTP-doc ont été réunies dans

rencontrées par la bibliothèque depuis 1997, dans ses différentes tentatives pour apprivoiser le numérique dans sa matérialité, ses usages et sa pérennité. Le changement d'échelle à la période suivante va précipiter le passage de l'expérimentation et la théorisation à la mise en pratique.

II. Se transformer politiquement et structurellement : industrialisation et professionnalisation (2003-2008)

Lorsqu'à l'automne 2004 Google annonce le programme « Google print », d'abord avec les éditeurs à la foire de Francfort, puis avec les bibliothèques en décembre⁴⁴, le domaine des bibliothèques numériques se trouve confronté à un changement d'échelle d'une ampleur telle qu'il va remettre en cause toutes les constructions de la période précédente, sur le plan théorique comme sur le plan technique. Bien sûr, le projet de la numérisation était dès l'origine un projet politique, adossé à l'annonce de la création d'une très grande bibliothèque par Mitterrand le 14 juillet 1988. Cependant, le tournant de l'année 2005 marque une étape décisive dans ce projet politique, largement provoquée par l'entrée en scène d'un nouvel acteur issu du privé et plus spécifiquement du web.

Jean-Noël Jeanneney, alors président de la BnF, porte le débat sur la numérisation et l'accessibilité en ligne du patrimoine sur la scène publique, en publiant d'abord une tribune dans *Le Monde*, le 23 janvier 2005, puis son ouvrage *Quand Google défie l'Europe*⁴⁵ en avril 2005. La période qui s'ensuit, sur laquelle nous ne reviendrons pas ici car elle est bien connue, correspond à la construction d'un projet politique national et européen, qui débouchera sur la mise à disposition de financements importants pour la numérisation des bibliothèques françaises par le Centre National du Livre (CNL), la création de la bibliothèque numérique européenne Europeana et la publication en 2010 du *Schéma numérique des bibliothèques*⁴⁶ présenté par Bruno Racine qui a alors succédé à Jean-Noël Jeanneney. L'enseignement supérieur se structure également autour du dispositif BSN (Bibliothèque Scientifique Numérique) créé en 2009, qui vise à faire émerger une dynamique au niveau national et à mobiliser des moyens, et dont l'un des enjeux réside dans l'acquisition de la documentation électronique au niveau national⁴⁷.

Historien des médias, Jean-Noël Jeanneney ne pouvait ignorer l'importance que revêtait l'archivage de l'Internet pour la construction de la collection patrimoniale de la BnF⁴⁸. La

44 Gaëlle Béquet, *op. cit.* p. 205.

45 Paris : Mille et une nuits, 2005.

46 En ligne : <https://www.ladocumentationfrancaise.fr/rapports-publics/104000143/>.

47 Cf. *Arabesques* n°68, nov-déc. 2012.

48 Il mentionne d'ailleurs cet enjeu dans son plaidoyer *Quand Google défie l'Europe*, p. 119 dans la 2^e édition datée de 2006.

période 2003-2010 est aussi celle où est créé le consortium IIPC⁴⁹ pour la préservation de l'Internet (2003), où voit le jour la loi française entérine un dépôt légal des contenus publiés en ligne, où la BnF lance ses premières collectes larges internalisées.

Nous proposons une lecture de cette période, dans la continuité de la précédente, au prisme de trois caractéristiques du web que la BnF s'approprie alors : la masse, la globalisation ou internationalisation, et la transversalité.

A. Changement d'échelle : gérer la masse

Le mythe de l'explosion documentaire, suivant lequel la qualité de la production s'effondre supposément sous le poids de la quantité au point de rendre la masse de documentation ingérable pour les professionnels en charge de l'organiser et la conserver, est si ancien qu'il est presque gênant de l'invoquer ici comme une caractéristique de l'ère numérique. En 1869, dans *Nos fils*, Jules Michelet se plaignait déjà d'être « *inondé de journaux, de romans, d'un déluge de papiers* ». La question du volume de données produites et échangées sur le web est néanmoins l'un des éléments clefs de compréhension des problématiques posées par le numérique à la BnF à cette période.

La masse est la caractéristique du web qui semble rendre vaine toute tentative d'en avoir une approche documentaire. Ainsi, dès l'origine, la notion d'exhaustivité inhérente au principe du dépôt légal tel qu'il était défini par le Code du Patrimoine a dû être abandonnée. Mais il a également fallu renoncer à l'idée de dépôt, puisque c'était désormais la BnF qui prenait l'initiative de la collecte et la réalisait à l'aide de robots : en lieu et place d'effectuer une veille pour relancer les déposants négligents, le bibliothécaire effectuait désormais une sélection, dont il fallut définir les critères pour que ceux-ci correspondent à l'esprit de représentativité du dépôt légal, bien différent des acquisitions⁵⁰. Le catalogage, enfin, était irréalisable à l'unité et devait faire place à un index automatique, ce qui remettait en cause l'idée de bibliographie nationale indissociable du dépôt légal. Ces renoncements, ces nouvelles stratégies étaient en place dès les premières collectes « larges » réalisées par Internet Archive pour la

49 International Internet Preservation Consortium.

50 Gildas Illien et Clément Oury, « Quelle politique documentaire pour l'archivage des sites Internet ? ». Dans Pierre Carbone et François Cavalier (dir.), *Les collections électroniques, une nouvelle politique documentaire*. Paris : Éditions du Cercle de la librairie, 2012, p. 157.

BnF en 2004⁵¹. En 2006, la loi sur les Droits d’Auteur et Droits Voisins dans la Société de l’Information (DADVSI) intègre l’archivage du web au cadre du dépôt légal, en étendant celui-ci à tous les services de communication au public en ligne⁵². Le décret qui précise ses modalités d’application, et entérine donc les caractéristiques particulières de ce dépôt légal « d’un genre entièrement nouveau », ne verra le jour que cinq ans plus tard⁵³.

Le déploiement de ce dispositif dans le début des années 2000 à la BnF montre que l’établissement avait pris la mesure du défi technique, organisationnel, économique et documentaire qu’était le web en tant que média à conserver. La logique de la masse fait son irruption dans le domaine jusqu’alors apparemment plus maîtrisable de la bibliothèque numérique en 2004, lorsque Google annonce son intention de numériser des bibliothèques entières. Nous ne reviendrons pas ici sur l’aspect politique de la prise de position de Jean-Noël Jeanneney, déjà évoquée. L’impact de la « numérisation de masse », qui débute en 2007, se traduit également dans le dispositif déployé par la bibliothèque pour gérer ce changement d’échelle dantesque⁵⁴. L’enjeu est de passer de 6.000 ouvrages numérisés par an à 100.000. Les moyens demandés en 2006 s’élèvent à 40 postes supplémentaires et 10 millions d’euros par an. Plus encore que l’archivage du web, la numérisation en masse des collections d’imprimés de la bibliothèque mobilise la direction des collections, qui doit assurer la sélection des ouvrages et la logistique.

Des questions se posent sur la conservation des ouvrages originaux et leur communication : un ouvrage en mauvais état de conservation peut-il être envoyé en numérisation ? Une fois l’ouvrage numérisé, doit-on limiter la consultation de l’exemplaire physique et renvoyer à la version en ligne, au titre de la conservation préventive ? Le mot d’ordre alors défini est qu’un ouvrage doit être communicable quand il part en numérisation, et le rester ensuite. Cependant, la numérisation apparaît désormais comme un moyen de reproduire les documents à

51 Les collectes étaient réalisées sous forme d’une prestation de service par la fondation Internet Archive, et les supports physiques sur lesquels les archives étaient enregistrées, les « Petabox », acheminées jusqu’à la BnF depuis les États-Unis. Les Petabox étaient à la fois un support de livraison, de stockage et d’accès. Cf. Sophie Derrot et Géraldine Camille, *Étude sur les priorités de versement des chaînes rétrospectives dans SPAR*, 2014. Archives numériques de la BnF : BnF-ADM-2014-091576.

52 Cf. **Emmanuelle Bermès, « Quand le dépôt légal devient numérique : épistémologie d’un nouvel objet patrimonial. » Dans *Quaderni : communication, technologies, pouvoir*, n°98, 2018-2019, p.73.

53 Décret n° 2011-1904 du 19 décembre 2011 relatif au dépôt légal. En ligne : <https://www.legifrance.gouv.fr/eli/decret/2011/12/19/2011-1904/jo/texte>.

54 Gaëlle Béquet, *op. cit.* p. 209-213.

fin de conservation comme l'avait été le microfilm : on parle enfin de numérisation de conservation⁵⁵.

L'exemple de Google conduit enfin à l'océrisation systématique des collections numérisées, ce qui pose un nouveau problème lié à la masse : celui de l'indexation de grandes quantités de texte non structuré. La BnF avait déjà commencé à développer ses compétences dans ce domaine : un premier moteur de recherche plein-texte existait dans la version de Gallica en ligne depuis 2000, et les techniques de collecte du web étaient inspirées de celles des moteurs de recherche généralistes. La recherche par mots de la notice dans le catalogue, utilisant le moteur de recherche *open source* Lucene, fut mise en service en 2009. Cependant, le moteur de Gallica 2000 n'indexait qu'une très petite quantité de texte, tandis que celui du catalogue travaillait sur une masse de données très structurées et homogènes. S'agissant des archives de l'Internet, la mise en place d'un moteur de recherche plein-texte était techniquement invivable⁵⁶. Le prototype Europeana, présenté au salon du livre de Paris en mars 2007, transformé l'année suivante en une nouvelle version de Gallica (« Gallica 2 ») est donc le premier essai de véritable indexation plein-texte de l'OCR brut en masse. Celle-ci est alors réalisée avec le moteur de recherche libre Lucene, qui sera remplacé en 2014 par l'outil commercial Exalead. La mise en place de ces nouveaux outils introduit les pratiques du web, notamment le classement des résultats par pertinence et l'affinage par facettes, dans l'univers de la bibliothèque, mobilisant de nouvelles compétences⁵⁷.

B. Internationalisation : la bibliothèque face à un média global

Une autre caractéristique de la période réside dans l'ouverture de la stratégie numérique de la bibliothèque au niveau national et international. La prise de position politique de Jean-Noël Jeanneney, qui vise d'entrée une réplique « européenne » à Google, repose en partie sur la conscience des moyens à mobiliser pour élaborer une réponse technologique à la hauteur de l'enjeu. Il appelle de ses vœux dans son *Plaidoyer* la mobilisation du génie industriel européen pour développer une expertise sur le traitement de données à grande échelle : c'est ce qui

55 Cf. **Emmanuelle Bermès, Marie-Élise Fréon et Frédéric Martin, « Tous les chemins mènent au numérique. ». Dans *Bulletin des bibliothèques de France*, 2008, n° 6, p. 34-39. En ligne : <http://bbf.enssib.fr/consulter/bbf-2008-06-0034-006>.

56 En 2008, Gildas Illien le définit comme « une priorité » tout en précisant qu'on n'a pas encore su lever les limitations techniques. « Le Dépôt légal de l'internet en pratique ». Dans *Bulletin des bibliothèques de France*, 2008, n° 6, p. 20-27. En ligne : <http://bbf.enssib.fr/consulter/bbf-2008-06-0020-004>.

57 Cf. **Emmanuelle Bermès, « Les Moteurs de recherche ». Dans *Bulletin des bibliothèques de France*, 2007, n° 6, p. 5-10. En ligne : <http://bbf.enssib.fr/consulter/bbf-2007-06-0005-001>.

débouchera sur le projet franco-allemand Quaero, visant à développer un moteur de recherche capable notamment d'indexer des contenus non textuels, dont la BnF est partie prenante⁵⁸. La même préoccupation conduit en 2003 à la création du consortium IIPC pour l'archivage de l'internet⁵⁹ : face à une activité entièrement nouvelle que l'ensemble des bibliothèques nationales ont à gérer de manière similaire, la mutualisation des outils, en commun avec la fondation Internet Archive qui avait joué un rôle de précurseur, semble s'imposer.

Cependant, au-delà de cet enjeu technologique, c'est encore la nature du web, par essence dépourvu de frontières géographiques ou politiques, qui impose sa loi et ses propres règles aux bibliothèques. Gildas Illien qualifie le questionnement sur l'archivage du web d'« *indéniablement politique* » du fait du hiatus qu'il crée entre l'identité et les missions des bibliothèques nationales, héritières du concept d'État-nation, et le caractère par nature global du réseau⁶⁰. La loi de 2006, en ajoutant au Code du Patrimoine un dépôt légal de l'Internet « français », crée en réalité un paradoxe aussi bien en termes de collecte que d'usage. Dans la vision idéale des fondateurs du consortium IIPC, notamment celle de Brewster Kahle qui dirige Internet Archive, l'objectif de cette alliance internationale était la constitution d'une archive mondiale et globale, que les frontières des domaines nationaux n'entraveraient pas. Mais devant l'impossibilité politique d'une centralisation qui aurait par trop remis en cause le cadre de la propriété intellectuelle, les bibliothèques se sont focalisées sur une « vision d'interopérabilité » consistant à partager les méthodes, visions, formats et outils qui faciliteraient potentiellement plus tard la création de liens entre les archives web nationales des différents pays.

Une telle finalité semblait en effet inéluctable au prisme des usages, globaux comme le web lui-même. Dans le domaine scientifique et documentaire, l'abolition des frontières et de la distance permise par le web a très tôt rendu évident et nécessaire le besoin de mettre fin à la logique de silos qui obligeait les internautes à consulter successivement et séparément chaque entrepôt de ressources potentiellement pertinent. Le protocole OAI-PMH⁶¹, conçu pour accéder de manière conjointe aux données des archives ouvertes, fut utilisé pour la première fois à la BnF dans le cadre du dossier thématique *La France en Amérique* (2005). Il était encore au cœur du prototype Europeana réalisé par la BnF pour démontrer la faisabilité de la biblio-

58 <http://www.quaero.org/>

59 Gildas Illien, « Une histoire politique de l'archivage du web ». Dans *Bulletin des bibliothèques de France*, 2011, n° 2, p. 60-68. En ligne : <http://bbf.enssib.fr/consulter/bbf-2011-02-0060-012>.

60 Gildas Illien, *ibid.*

61 Open Archive Initiative Protocol for Metadata Harvesting.

thèque numérique européenne et présenté au salon du livre de Paris en 2007⁶², agrégeant trois bibliothèques : Gallica, la bibliothèque nationale de Hongrie et celle du Portugal. En 2008, le nouveau portail européen était lancé sur la même base technique⁶³, et la nouvelle version de Gallica devenait une bibliothèque collective, intégrant les métadonnées de bibliothèques partenaires. La politique de coopération numérique de la BnF au sein d'un réseau national était appelée à prendre une importance considérable.

Le protocole OAI-PMH dut son succès initial, non démenti à ce jour, au fait qu'il s'ancrait dans les pratiques du web, en rupture avec une informatique plus lourde utilisée dans les bibliothèques⁶⁴. En outre, l'espoir qu'il serait adopté par d'autres acteurs du web, comme les éditeurs⁶⁵ mais surtout les moteurs de recherche, permettait d'imaginer qu'il deviendrait un outil pour améliorer la visibilité des bibliothèques numériques sur le web à travers leur référencement par les moteurs de recherche comme Google⁶⁶. L'enjeu était d'accroître les flux d'audience par le partage des métadonnées tout en gardant la main sur la consultation des documents numérisés eux-mêmes pour continuer à comptabiliser les visites. D'autres leviers comme la mise en place des identifiants pérennes ARK, déployés en 2005 pour Gallica et peu après pour le catalogue général, servirent toutefois cet objectif de référencement avec plus de succès.

C. Transversalité : quand le web bouscule l'organisation

Dans son *Histoire des médias*⁶⁷, Jean-Noël Jeanneney compare la structure du web à un « diabololo » ou à deux entonnoirs placés bout à bout, reliés par leur embout le plus étroit : un grand nombre de producteurs s'y adressent à un grand nombre d'interlocuteurs. Si Tim Berners-Lee avait pensé dès l'origine le web comme un média lisible et inscriptible, le « *read-write web* »⁶⁸ ne devient une réalité qu'au début des années 2000, au moment où les outils du

62 G. Béquet, *op. cit.* p. 209.

63 Cf. **Emmanuelle Bermès et Elisabeth Freyre, « Une contribution française à la bibliothèque numérique européenne : Europeana et la Bibliothèque nationale de France ». Dans *Digitalia*, vol. 3 n°1, 2008.

64 Notamment le format MARC et le protocole Z39.50. Cf. Yves Desrichards, *Cinquante ans de numérique en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2017, p. 65.

65 Dans le cadre du projet d'intégration d'œuvres sous droit à Gallica : cf. G. Béquet p.215 et article dans la revue *Digitalia*, *op.cit.*

66 Clément Oury, *Le département de la bibliothèque numérique à la Bibliothèque nationale de France*. Rapport de stage. Villeurbanne : ENSSIB, 2005, p. 10. En ligne : <http://enssibal.enssib.fr/bibliotheque/documents/dcb/rsoury.pdf>

67 Paris : Seuil, 2011.

68 Dan Gillmor, *We the media: grassroots journalism by the people, for the people*. Sebastopol : O'Reilly, 2006.

web dit « 2.0 » – blogs, réseaux sociaux etc. – permettent au « grand public » de produire et mettre en ligne des contenus en un instant, gratuitement et sans compétences particulières. Dans la bibliothèque numérique, l’usager devient soudain un acteur et plus seulement un lecteur. La création dans Gallica 2 d’un espace personnel ainsi que la réflexion sur les fonctionnalités de groupes et de partages participèrent de l’appropriation par la BnF de ces enjeux. Cependant, les retours négatifs des premières études d’usages sur Gallica⁶⁹ conduisirent dans un premier temps la BnF à une attitude prudente à l’égard de ces initiatives collaboratives.

Dans le numérique, la transversalité se manifeste également par une dépendance accrue entre les différentes opérations successivement réalisées, qu’il s’agisse de numérisation ou d’autres projets numériques. Le projet de numérisation de la presse quotidienne nationale voulu par Jean-Noël Jeanneney et annoncé par la BnF en février 2005, qui fut le premier à bénéficier systématiquement d’un OCR, impliqua un questionnement sur tous les maillons de la chaîne, des usages spécifiques de ce type de document – qui firent l’objet d’une étude – à la numérisation en passant par la structuration des états de collection dans le catalogue⁷⁰. L’interdépendance entre les choix techniques réalisés à différentes étapes justifiait une vision transverse et de nombreuses interactions entre les acteurs. La relation entre informaticiens et bibliothécaires, en particulier, était profondément transformée.

Plus tard c’est la problématique de la masse, comme évoqué plus haut, qui conduisit à déployer à l’échelle de la bibliothèque une organisation complexe, impliquant notamment tous les départements de collection, qui avaient la charge d’alimenter en continu les ateliers de numérisation pour produire les cent mille ouvrages par an attendus. Le même type d’organisation transverse, reposant sur des réseaux de correspondants par discipline, fut mis en place pour assurer la sélection des sites web des collectes dites « ciblées » des archives de l’Internet⁷¹.

Le département de la bibliothèque numérique se réorganisa progressivement à cette période, se délestant des activités qui pouvaient être rattachées à des processus classiques de la

69 Philippe Chevallier, *Résumé des précédentes études d’usage Gallica*. 6 mars 2013. Archives numériques de la BnF : BnF-ADM-2009-084330-01.

70 Benjamin Prémel, « La numérisation de la presse rétrospective à la BnF ». Dans *Le Temps des médias*, 2007/1 (n° 8), p. 267 à 278. En ligne : <https://www.cairn.info/revue-le-temps-des-medias-2007-1-page-267.htm>

71 Gildas Illien, « Le Dépôt légal de l’internet en pratique », *op. cit.*

bibliothèque⁷² pour se concentrer sur celles qui, plus innovantes, exigeaient encore une dimension exploratoire. Celles-ci furent réparties en deux pôles : d'une part celui qui traitait de l'archivage du web, et d'autre part celui qui poursuivait le travail technique et fonctionnel de stabilisation de la bibliothèque numérique, notamment à travers les métadonnées. Lorsqu'en 2007, au départ de Catherine Lupovici sa directrice, il fut question de supprimer ce département qui n'avait plus la taille critique, plusieurs scénarios furent envisagés. Finalement, le rattachement de l'archivage du web au département du dépôt légal entérina sa nature de processus industrialisé, visant un « support » indépendant des disciplines, comme le livre. Le pôle métadonnées fut intégré à l'ancienne agence bibliographique nationale, au profit d'une réorganisation qui la transforma en département de l'information bibliographique et numérique. Enfin, pour tenter de prolonger la dynamique de transversalité que le département avait su insuffler dans l'établissement autour des projets numériques, fut créé l'observatoire ORH-ION (Observatoire des Ressources Humaines sous l'Impact Opérationnel du Numérique), première communauté de pratique informelle et non hiérarchique, dont l'invention était une émanation directe de l'observation de la structuration spontanée des communautés en ligne⁷³. Ces transformations organisationnelles s'achevèrent en 2008 avec la suppression du département de la bibliothèque numérique, soutenue par l'affirmation que désormais, toute la bibliothèque était devenue numérique.

Aux caractéristiques du web que nous avons étudiées dans les lignes qui précèdent, il faut ajouter en 2008 une dimension fondamentale, qui s'est progressivement renforcée avec l'adoption à grande échelle des terminaux mobiles (dont les smartphones et tablettes) et la disponibilité croissante et omniprésente du haut débit par wifi et réseau cellulaire : la convergence. Le numérique absorbe désormais de manière indifférenciée toutes les formes de publication qui existaient auparavant sur différents supports. Textes, images, sons, vidéos et objets multimédias interactifs se retrouvent ainsi en ligne ; il en est de même pour les différentes formes de publication, de la presse aux jeux vidéo en passant par les articles scientifiques, les

72 La numérisation a ainsi rejoint la conservation, les informaticiens ont été rattachés au DSI, le service de communication Internet à la délégation à la communication... Cf. **Emmanuelle Bermès et Louise Fauduet, « The Human Face of Digital Preservation: Organizational and Staff Challenges, and Initiatives at the Bibliothèque nationale de France. » Dans *The International Journal of Digital Curation* vol. 60 n°1, 2011, pp. 226-237.

73 *Ibid.*

livres, les encyclopédies, la musique, les logiciels, les cartes, les partitions, etc. Une telle convergence interroge nécessairement, pour la BnF, la notion de collection.

En 2001⁷⁴, après avoir détaillé les collections réparties dans les départements thématiques et celles qui s'organisent par support, Jacqueline Sanson, alors directrice des collections, rappelle :

Quelle que soit leur nature, quelle que soit leur valeur, quels que soient leur mode ou leur date d'entrée, quel que soit leur département d'affectation, les collections patrimoniales de la Bibliothèque nationale de France forment un tout : elles ne peuvent s'appréhender que dans leur unité.

À l'époque, cette unité des collections n'englobe pas le numérique, qui n'est considéré que sur le plan des services aux publics. En 2008, sept ans plus tard, les processus et les outils du numérique sont en place, faisant émerger une masse de nouveaux objets à produire, collecter, conserver et diffuser. La convergence propre au web, adossée à la transformation numérique de l'établissement, va désormais poser la question du statut de ces nouveaux objets, en continuité ou en rupture avec les collections de la bibliothèque.

74 Daniel Renoult et Jacqueline Melet-Sanson (dir.), *La Bibliothèque nationale de France : collections, services, publics*. Paris : Éditions du Cercle de la librairie, 2001, p. 73.

Deuxième partie :

Collecter, conserver, diffuser :
l'émergence des collections numériques (2008-2018)

Numériser les 11 millions d'ouvrages qui faisaient déborder les réserves de l'ancienne bibliothèque nationale de la rue de Richelieu : telle était l'ambition de la bibliothèque « d'un genre entièrement nouveau » qu'avait rêvée Jacques Attali et qu'avait commandée François Mitterrand. En 1988, en pleine époque du Minitel, il n'était pas aberrant d'imaginer la création d'un réseau informatique spécifique, dont la nouvelle bibliothèque serait la tête, et qui permettrait de distribuer jusque dans les collèges et lycées de toute la France ce trésor de la connaissance désormais à la portée de tous. Sans doute l'utopie de la reproduction universelle, pointée par Frédéric Barbier⁷⁵, jouait-elle ici à plein : nul ne songea à se demander s'il était possible ou même souhaitable de ressusciter ces témoins endormis de siècles d'histoire, certes, mais aussi de fanatisme religieux, de colonialisme, d'antisémitisme, de pornographie, d'erreurs scientifiques, ou tout simplement de répétitions ordinaires et sans grand intérêt. Combien faut-il numériser de missels du XIX^e siècle pour témoigner de la piété de la société à cette époque, et quel budget faut-il y consacrer ? Pour quel public – combien de chercheurs studieux consacreront leur thèse à ce matériau, pour le renvoyer ensuite à son sommeil, sous forme d'octets sur un serveur désormais ? On ne songea pas non plus à ce qu'il adviendrait de ce qui, dans les collections de la bibliothèque, n'était pas du texte : les cartes, les globes, les estampes, les photographies, les monnaies romaines, le trône de Dagobert, les maquettes de décors de l'Opéra de Paris, les 78 tours, les phonogrammes de Charles Cros... L'existence d'un programme parallèle de numérisation d'images, dans une logique de corpus iconographique, ne répondait que très partiellement à la question de la place de ces collections dans l'évolution de la Bibliothèque nationale vers son nouvel état numérique⁷⁶.

L'idée de la « bibliothèque d'un genre entièrement nouveau » était donc une utopie : celle d'une nouvelle collection maîtrisable, homogène, domestiquée ; une collection dont tous les objets seraient à un niveau d'égalité face à la soif républicaine de connaissance et qu'on porterait en conséquence à un égal niveau d'accessibilité, grâce au numérique. Une nouvelle collection dont les qualités lui permettraient de dépasser, voire de remplacer purement et simplement l'ancienne.

75 Barbier, Frédéric. « Patrimoine, production, reproduction ». Dans *Bulletin des bibliothèques de France*, 2004, n° 5, p. 11-20. En ligne : <http://bbf.enssib.fr/consulter/bbf-2004-05-0011-002>.

76 Cf. Catherine Éloi, « Les images fixes numérisées à la Bibliothèque nationale de France ». Dans *Bulletin des bibliothèques de France*, 2001, n° 5, p. 80-86. En ligne : <http://bbf.enssib.fr/consulter/bbf-2001-05-0080-011>.

Pour qui connaît la diversité des collections de la BnF, un tel projet paraît à la fois surdimensionné et terriblement insuffisant. Surdimensionné, parce que l'utopie d'une numérisation complète des fonds d'imprimés est inatteignable au regard des réalités économiques, techniques, juridiques et scientifiques. Insuffisant parce que les caractéristiques du web, que la bibliothèque s'est appropriées au cours de la décennie qui a précédé, dépassent largement l'idée d'une bibliothèque numérique centrée sur le livre et uniquement préoccupée de transférer les connaissances de l'ancien monde (ses collections d'imprimés) vers le nouveau.

De 1997 à 2007, la bibliothèque a expérimenté son rapport au numérique, ce nouvel environnement qui à la fois lui ressemble et lui est radicalement opposé ; dans la décennie suivante elle y plonge à pieds joints, en affrontant la masse et en démultipliant les moyens qu'elle lui consacre. Notre propos est ici de montrer comment s'est élaborée la notion de collection numérique, en lien avec les opérations de gestion réalisées par les professionnels d'une part, et avec les usages de recherche d'autre part.

I. Peut-on parler de collection numérique ? Les premières pierres d'une légitimation (2008-2011)

En 2010⁷⁷, nous nous interrogeons sur la validité du concept de collection numérique, tout en observant que « *nier la validité de ce concept reviendrait à mettre en cause l'idée même d'une présence spécifique et forte des bibliothèques sur Internet, tant la collection est inhérente aux bibliothèques, dont elle structure et conditionne l'ensemble de l'activité.* » Alors que la notion de document numérique avait, nous l'avons vu, suscité de nombreuses et profondes interrogations, celle de collection numérique restait inexplorée. Pourtant, vue des bibliothèques, l'idée de collection ne se contente pas d'englober la pluralité des documents : elle les fonde. Redécouvert dans les années 2000 à la lumière des interrogations sur le numérique, le texte de Suzanne Briet *Qu'est-ce que la documentation* écrit en 1951⁷⁸ propose ainsi, de façon audacieuse, l'idée que tout objet ou même tout être vivant peut être considéré comme un document, dès lors qu'il fait l'objet d'un travail de catalogage, de conservation, de documentation et de valorisation dans le cadre d'un ensemble plus vaste et organisé dont il devient l'un des éléments constitutifs :

Une étoile est-elle un document ? Un galet roulé par un torrent est-il un document ? Un animal vivant est-il un document ? Non. Mais sont des documents les photographies et les catalogues d'étoiles, les pierres d'un musée de minéralogie, les animaux catalogués et exposés dans un zoo.

En somme, la proposition de Suzanne Briet, qui sera reprise en 2003 dans le texte de Roger T. Pédaque⁷⁹, revient à conditionner le statut de document à la réalisation d'un certain nombre d'opérations de gestion correspondant aux missions et activités de la bibliothèque. C'est bien la collection qui fonde la reconnaissance du document en tant que tel.

Collecter, signaler, conserver, communiquer, diffuser : ces verbes qui décrivent les missions fondamentales de la bibliothèque sont utilisés en 2008 pour structurer le premier *Sché-*

77 Cf. **Emmanuelle Bermès et Frédéric Martin, « Le Concept de collection numérique ». Dans *Bulletin des bibliothèques de France*, 2010, n° 3, p. 13-17. En ligne : <http://bbf.enssib.fr/consulter/bbf-2010-03-0013-002>.

78 Texte revu et mis en ligne en juin 2008 par Laurent Martinet. En ligne : <http://martinetl.free.fr/suzannebriet/questcequeladocumentation/>.

79 Roger T. Pédaque, *Document : forme, signe et médium, les re-formulations du numérique*. 2003. En ligne : https://archivesic.ccsd.cnrs.fr/sic_00000511/document.

ma numérique de la BnF, une carte heuristique qui donne une vue d'ensemble de tous les projets numériques de l'établissement. S'y ajoutent « coopérer » et « organiser » qui reflètent l'impact du numérique sur l'organisation interne de la bibliothèque et son articulation avec son réseau national et international. Pour la première fois, la grille de lecture proposée pour les projets numériques épouse donc au plus près l'ensemble des missions de la bibliothèque, dans toutes leurs dimensions. Dans la période qui suit, la bibliothèque va effectuer un travail de réalignement entre les bouleversements apportés par le web et ses missions fondamentales, conduisant *in fine* à la légitimation du numérique en tant que collection.

A. Renouer avec la gestion des collections numériques

En 2002, François Stasse, qui vient de quitter ses fonctions de directeur général de la BnF, diagnostique dans *La véritable histoire de la grande bibliothèque*⁸⁰ les causes de ce qu'il considère comme l'échec du grand projet d'Attali. Il pointe alors deux obstacles intrinsèques au projet de numérisation : d'une part, le cadre juridique de la propriété intellectuelle qui interdit de reproduire une grande partie des ouvrages au mépris des droits (et des revenus) des auteurs et des éditeurs, et d'autre part, l'incertitude quant à la capacité à conserver de manière pérenne les documents numérisés. Comment pourrait-on considérer un ensemble de documents que l'on ne peut ni vraiment choisir, ni être sûr de préserver, comme une véritable collection ? Cette interrogation existentielle traverse les premières années de l'expansion du numérique à la BnF.

Tout d'abord, l'idée de politique documentaire, fondatrice d'une approche des collections propre au métier de bibliothécaire qui repose sur une sélection raisonnée, est bousculée par le numérique. Pierre Carbone écrit en 2009, dans son introduction au livre qu'il dirige avec François Cavalier⁸¹ :

Les professionnels des bibliothèques (...) ont de plus en plus souvent l'impression de ne plus effectuer un travail de sélection des acquisitions en fonction de leur évaluation des besoins de leurs publics, mais de promouvoir des offres éditoriales. La notion de collection, qui exprime cette sélection de l'offre et la constitution d'un ensemble, (...) ne s'est pas encore imposée pour l'électronique.

80 F. Stasse, *op. cit.*

81 *Les collections électroniques, une nouvelle politique documentaire*. Paris : Éditions du Cercle de la librairie, 2009, p. 9.

L'obligation d'acheter en bloc auprès des grands éditeurs numériques des bouquets de ressources déjà constitués prive les bibliothécaires de leur marge de manœuvre dans ce domaine. Il faut y ajouter l'incertitude quant à la pérennité de ces ressources pour la bibliothèque, puisqu'elles font l'objet d'abonnements qui supposent la fin de l'accès si la bibliothèque n'a plus les moyens de les maintenir. Les grands projets numériques des années 2006-2007, qu'il s'agisse de la numérisation de masse ou de l'archivage du web, posent le même type de problème : lorsqu'on numérise de manière systématique toute une liste de cotes dans l'ordre séquentiel⁸² ou que l'on collecte tous les sites en « .fr », l'idée de politique documentaire semble voler en éclat.

La relation ambiguë des objets issus du web au catalogue peut également contribuer à expliquer une réticence initiale à les considérer comme des collections. S'il est vrai que le catalogue fonde la collection en validant l'existence des objets qui la constituent et en permettant l'accès⁸³, les documents numériques peuvent faire figure de citoyens de seconde zone. Les documents numérisés, assimilés à de simples reproductions, n'apparaissent dans le catalogue que comme des exemplaires, rattachés à des notices bibliographiques qui ne sont pas corrigées à l'occasion de l'envoi en numérisation. Les archives de l'Internet ne font pas l'objet d'une bibliographie nationale, ce que valide le décret de 2011, et ne sont donc pas cataloguées. Même les ressources électroniques acquises, malgré plusieurs tentatives, restent à cette époque exclues du catalogue en raison de la volatilité des accès par licences sur les plateformes des éditeurs⁸⁴. Seuls les documents numériques sur support, qui entrent au département de l'audiovisuel, font l'objet d'un traitement bibliographique dans le catalogue général, au prix d'adaptations parfois périlleuses⁸⁵.

La spécificité des documents numériques, telle qu'elle s'est stabilisée dans la pratique professionnelle depuis dix ans, a donc conduit à les écarter des activités de traitement des collections classiques. Cependant, le changement d'échelle de 2006-2007, en imposant une ré-

82 La numérisation de masse repose théoriquement sur l'idée de numérisation systématique de magasins entiers, sans faire de tri car celui-ci s'avérerait plus coûteux. À la BnF, la cotation Clément permet d'identifier des segments importants de collection correspondant à un domaine ou un thème, ce qui facilite cette approche systématique.

83 Gautier Poupeau, « Histoire(s) de notices. » Dans Lisette Calderan, Pascale Laurent, Hélène Lowinger et Jacques Millet (dir.), *Le document numérique à l'heure du web*. Paris : ADBS, 2012, p.25-40. En ligne : <https://hal.inria.fr/hal-00740295v2>.

84 *Numérique horizon 2015 : Synthèse des orientations*. Archives numériques de la BnF : BnF-ADM-2014-101054.

85 Pour le jeu vidéo, Benjamin Barbier aborde spécifiquement cette question dans la deuxième partie de sa thèse, *op. cit.*, chapitre 5, p. 116.

partition des tâches de sélection sur l'ensemble des départements de collection, aussi bien pour les archives du web que pour la numérisation dite désormais « de masse », force l'institution à réintroduire des règles et pratiques de gestion. Ces opérations ne sont en réalité pas aussi exhaustives qu'elles le prétendent, et impliquent des critères de sélection. Ceux-ci peuvent être d'ordre technique, de la « profondeur » des sites web qui se mesure en nombre de « clics » depuis la page d'accueil à « l'ouvrabilité » des volumes reliés qui doivent passer sous les scanners. D'autres portent sur le contenu des documents et sont dictés par les nouvelles conditions de constitution des collections : représentativité pour les sites web qui sont sélectionnés au titre d'un échantillonnage de la production nationale, date de la mort du ou des auteurs qui déterminent la situation juridique d'un ouvrage à numériser.

Le temps de travail important consacré à ce travail de sélection, qu'il soit intellectuel, juridique, technique ou autre, en vient à justifier l'idée que le résultat obtenu ne pourrait être facilement reconstitué, qu'il est donc précieux et doit être conservé. Il débouche sur la création de données de gestion qui, si elles ne sont pas stockées dans le catalogue, présentent un intérêt certain pour la connaissance de la collection numérique autant que pour son élaboration. Ainsi l'application BCweb, utilisée par les chargés de collection pour renseigner les sites à collecter et les paramètres de collecte (profondeur, fréquence...), est-elle une ressource jugée comme fondamentale par les chercheurs qui ont besoin de connaître les conditions de création de l'archive pour comprendre les corpus qu'ils exploitent⁸⁶. Si on ne peut parler de catalogage, il existe bien des données liées à la sélection, donc au traitement des collections, dont il faut déterminer la structure, l'usage et la pérennité.

Le cas de l'application de sélection pour la numérisation, dite « AD-CAT-15 », illustre particulièrement bien cette prise de conscience. En 2007, l'urgence est de produire en masse pour atteindre les 100.000 documents par an qui conditionnent le financement de la numérisation par le Centre National du Livre. Les agents de la direction des collections, mobilisés en nombre, ont besoin d'un outil qui leur permette d'établir, à partir d'extractions du catalogue, des listes d'ouvrages propres à être numérisés. Plus précisément, il s'agit d'exclure des listes les ouvrages impropres à la numérisation, puisque le principe de la numérisation de masse est de renoncer à une sélection intellectuelle au profit de la reproduction de segments entiers de magasins, à l'exception des quelques volumes faisant l'objet de critères d'exclusion principa-

86 Valérie Baudouin et Zeynep Pelhivan, « La mise en place d'un dispositif méthodologique » dans Valérie Baudouin, Philippe Chevallier et Lionel Maurel (dir.) *Le web français de la grande guerre : réseaux amateurs et institutionnels*. Nanterre : Presses universitaires de Paris Nanterre, 2018, p. 117-118.

lement techniques. Développée dans l'urgence, l'application AD-CAT-15, peu robuste, est censée être utilisée de manière ponctuelle dans un processus de numérisation linéaire : création d'une liste, exclusion des volumes non numérisables, envoi en numérisation, réception des fichiers, mise en ligne. Il est donc prévu qu'elle puisse être régulièrement purgée de ses données. Quelques années plus tard, lorsque la question de cette purge est à nouveau posée, on réalisera que cette application est le seul endroit où sont renseignées les raisons pour lesquelles un livre a été exclu de la numérisation à un moment donné : ouvrage sous droit, problème d'état de conservation ou d'ouvrabilité, typographie impropre à l'OCR... Or, si les règles qui régissent les programmes de numérisation évoluent et s'assouplissent, il n'est pas question de refaire ce travail colossal d'analyse. Cette application de gestion, qui fait le lien entre la collection physique et la collection numérique, est devenue indispensable. À l'heure où nous écrivons ces lignes, elle est toujours en service.

B. Retrouver l'unité des collections grâce à la préservation numérique

L'autre faille identifiée par Stasse en 2002, celle de la préservation du numérique, fait l'objet dans les mêmes années d'un effort considérable. En 2006, lorsque l'ADBU organise sa journée d'études annuelle sur le thème de la préservation des documents numériques, c'est Catherine Lupovici, alors directrice du département de la bibliothèque numérique, qui ouvre la séance⁸⁷. Sa préoccupation pour cet enjeu n'était pas nouvelle : la légende raconte que lors de sa première réunion de service à la BnF, elle aurait jeté à la poubelle un CD-ROM comme ceux qui étaient alors utilisés pour stocker la numérisation, pour marquer les esprits par ce geste irrévérencieux. Au prix d'un effort de sensibilisation constant pendant ces années, l'idée d'adopter la norme OAIS⁸⁸ issue du domaine de l'aérospatiale parvient à s'imposer⁸⁹. L'instruction du futur système SPAR (Système de préservation et d'archivage réparti), qui débute à cette époque, fait ainsi émerger un nouveau domaine de gestion, celui de la conservation, contribuant de ce fait à renforcer l'idée que ce qui est constitué est bien une collection.

87 Cavalier, François. « La Préservation des documents numériques ». Dans *Bulletin des bibliothèques de France*, 2002, n° 1, p. 97-99. En ligne : <http://bbf.enssib.fr/consulter/bbf-2002-01-0097-003>.

88 Open Archival Information System, norme publiée sous la référence ISO 14721 en 2003.

89 Catherine Lupovici, « Gestion de la pérennisation des documents numériques : aspects fonctionnels et techniques. » Dans Lisette Calderan, Bernard Hidoine et Jacques Millet (dir.), *Pérenniser le document numérique*. Séminaire INRIA, 2-6 octobre 2006, Amboise. Paris : ADBS, 2006, p. 51-82.

La convergence de la communauté des bibliothèques et des archives avec celle de la conquête spatiale – à travers son centre de documentation, le CCSDS⁹⁰ – a déjà en soi quelque chose d’historique. À l’échelle de la bibliothèque, on peut en dire autant de l’impact du projet SPAR sur la conception de la notion de collection numérique. En effet, dès l’origine, il est envisagé que ce système unique assure la pérennisation de l’ensemble des ressources numériques de la BnF, sans considération de l’organisation de la bibliothèque. En plus de la numérisation et des archives de l’Internet, son périmètre inclut en théorie les acquisitions d’archives de ressources électroniques pérennes réalisées auprès d’éditeurs comme Springer, Elsevier et Wiley-Blackwell⁹¹, la production documentaire de l’établissement (bases Lotus Notes notamment), d’autres filières de dépôt légal dématérialisé et pour la première fois, les documents audiovisuels et multimédias. Dans le cadre de la coopération numérique, SPAR est également appelé à préserver les collections de bibliothèques partenaires numérisées grâce au financement du CNL, constituant ainsi une « collection numérique nationale ».

Il ne faut pas minimiser l’impact de l’intégration, même théorique, de l’audiovisuel dans SPAR. La gestion des collections vidéos, sonores et multimédias était depuis l’ouverture du site François-Mitterrand opérée dans un système propre, le système audiovisuel (SA), qui avait été développé en parallèle du système d’information (SI). Rares sont ceux qui se souviennent que le poste de lecture audiovisuel avait ouvert au public dès 1996 avec succès, alors que deux ans plus tard l’informatique défaillante de la « très grande bibliothèque » toute neuve était pointée du doigt par les médias⁹². Le SA disposait de son propre système de stockage, de sa propre équipe informatique, de son interface d’accès, de ses postes informatiques. Seul le catalogue fut intégré à celui des autres documents, alors nommé BN-Opale, qui devint à cette occasion BN-Opale Plus.

En prévoyant à terme d’intégrer les collections audiovisuelles, SPAR posait la première pierre de la future convergence entre SI et SA, qui sera décidée par le schéma directeur informatique en 2017⁹³. Mais surtout, il posait un cadre théorique et organisationnel qui faisait le lien entre les pratiques de gestion de l’ancien monde, celui où les documents – même numériques – étaient sur des supports, et celui du web. Ainsi le « plan de sauvegarde audiovisuel », qui depuis son origine mobilisait la reproduction numérique comme technique de conserva-

90 Consultative Committee for Space Data Systems.

91 Michel Netzer, *Les sciences en bibliothèque*. Paris : Éditions du cercle de la librairie, 2017, p. 173.

92 François Stasse, *op. cit.* p. 143-144, mentionne la débâcle de l’informatique sans dire un mot du SA.

93 Cf. archives numériques de la BnF, BnF-ADM-2017-119867 : *Schéma directeur des systèmes d’information. Synthèse au 31 mars 2017*.

tion, contribua-t-il à légitimer l'idée de numérisation de conservation. L'émulation des environnements informatiques des années 1980, qui avait permis de faire revivre les jeux éducatifs sur TO7⁹⁴, apporta des pistes de réflexion pour la pérennisation des archives du web. Dans SPAR, tous ces documents seraient à terme dématérialisés, documentés, et gérés de façon globale via l'articulation des savoir-faire des bibliothécaires sur les collections avec les compétences techniques des informaticiens.

Le développement du système commence en 2008 et il est mis en production en 2010⁹⁵. Si toutes les filières, correspondant aux différents types de documents, ne furent développées que progressivement entre 2008 et 2014, cette intention première – que l'on peut assimiler à une approche plus archivistique que bibliothéconomique – témoigne de la reconstitution d'une unité de la collection qui dépasse les clivages traditionnels entre documents numérisés et nés-numériques, ou entre documents dématérialisés et sur support. Si on appelle alors parfois ce système « la cinquième tour »⁹⁶, c'est pour pointer la continuité entre les réserves qui abritent les livres et ce nouveau « magasin numérique » qui exigera, au même titre que des collections plus traditionnelles, une véritable politique de conservation, qui ne soit pas pilotée uniquement par l'informatique, mais par les professionnels des bibliothèques.

C. Repenser la visibilité de la bibliothèque sur le web

Pour être reconnue comme telle, une collection de bibliothèque doit être non seulement signalée et conservée, mais également diffusée et valorisée. La question de la visibilité sur le web de la BnF, après plus de dix ans de construction empirique, se pose avec acuité : la dispersion des ressources mises en ligne, leur foisonnement et la multiplication des bases de données et des sites rendent difficilement intelligible, pour une personne extérieure à la bibliothèque, toute l'étendue des contenus et services disponibles. Si un lecteur s'intéresse à Victor Hugo, comment lui indiquer qu'il trouvera ses ouvrages dans le catalogue général, ses manuscrits autographes dans le catalogue Archives et manuscrits, d'anciennes éditions numérisées de ses romans dans Gallica, une exposition virtuelle qui lui est dédiée, etc. ? C'est ainsi

94 Alain Carou et Jean-Philippe Humblot, *Retour d'expérience sur l'émulation à la BnF*. Présentation pour la réunion du groupe Pérennisation de l'Information Numérique (PIN), 10 décembre 2004. En ligne : http://pin.association-aristote.fr/doku.php/public/reunion_pleni%C3%A9re/cr/cr_20041210.

95 Cf. ** Emmanuelle Bermès, Isabelle Dussert-Carbone, Thomas Ledoux et Christian Lupovici, « *Digital preservation at the National Library of France: a technical and organizational overview* ». IFLA WLIC 2008 : IFLA World Library and Information Congress, 74th IFLA General Conference and Assembly. Montréal, 2008. En ligne : http://archive.ifla.org/IV/ifla74/papers/084-Bermes_Carbone_Ledoux_Lupovici-en.pdf

96 « La boîte à outils de la numérisation. » Dans *Chroniques*, juin 2007. En ligne : <https://c.bnf.fr/D1E>.

que naît en 2008 l'idée d'un « pivot documentaire », un site qui agrègerait toutes les informations et ressources de la bibliothèque autour des auteurs et des œuvres qui sont au cœur des collections⁹⁷.

Derrière ce projet, on retrouve l'utopie de la bibliothèque « d'un genre entièrement nouveau » – une source organisée de connaissance, accessible en ligne, centrée sur les grands textes de la culture française et européenne – mais aussi, d'une certaine manière, l'héritage de son projet politique. Après avoir voulu contrer Google avec Jean-Noël Jeanneney, la bibliothèque adopte sous la présidence de Bruno Racine une approche « raisonnée » qui prend acte des apports du géant du web à la société et à la diffusion de la culture, et imagine des stratégies de coopération et de complémentarité⁹⁸. Parmi celles-ci, la question du référencement est clef : il est désormais acquis que plus de 90% des internautes utilisent Google comme point d'entrée de leurs recherches, voire de toute session de navigation sur le web. Pour rester pertinente dans un monde numérique, la bibliothèque doit donc s'efforcer de faire apparaître ses ressources le plus haut possible dans la page de résultats du moteur de recherche généraliste.

Le site finalement nommé Data.bnf.fr et lancé en 2011 conjugue cette ambition, à travers la proposition de pages simples, bien structurées et riches en liens, avec une expérimentation des technologies du web sémantique⁹⁹. Son objectif : jouer un rôle de « hub » organisant les ressources autour des œuvres et des personnes qui sont au centre de l'intérêt des internautes, assurant une bonne visibilité dans les moteurs de recherche, et orientant les internautes vers Gallica et les autres ressources en ligne de la BnF. Au cœur du projet Data.bnf.fr, mais aussi de la transition bibliographique vers des modèles « orientés entités »¹⁰⁰ qui est en train de voir le jour au niveau international, on trouve une idée clef : il est devenu plus important de répondre aux questions que se posent les internautes (qui ? quoi ? où ? quand ?) que de proposer des listes de documents. Lorsque Google lance le Knowledge Graph¹⁰¹ en 2012, il confirme la validité de cette stratégie : alors qu'auparavant, face à une question posée, Google

97 *Étude marketing pour la préfiguration du pivot documentaire*, 2010. Archives numériques de la BnF : BnF-ADM-2010-007185.

98 B. Racine, *Google et le nouveau monde*. Paris : Plon, 2010.

99 Cf. **Emmanuelle Bermès, Vincent Boulet et Céline Leclaire, « Améliorer l'accès aux données des bibliothèques sur le web : l'exemple de data.bnf.fr ». IFLA WLIC 2016 : IFLA World Library and Information Congress, 82nd IFLA General Conference and Assembly. Columbus, 2016. En ligne : <http://library.ifla.org/1447/>.

100 Ce sujet a été au cœur de la conférence satellite de l'IFLA organisée à la BnF en 2014 sur le thème : *Linked Data in Libraries: Let's make it happen!* dont les actes sont en ligne : <http://ifla2014-satdata.bnf.fr/program.html>. Un groupe d'intérêt spécialisé sur le web sémantique a pu ensuite être créé au sein de l'IFLA.

101 Amit Singhal, « Introducing the Knowledge Graph: things, not strings ». *Google blog*, 16 mai 2012. En ligne : <https://www.blog.google/products/search/introducing-knowledge-graph-things-not/>.

répondait par une liste de documents – des pages web – désormais il propose une fiche signalétique qui cherche à répondre directement au besoin informationnel à partir de données structurées. Avec ce projet, la bibliothèque tente un nouveau repositionnement par rapport au web : il ne s’agit plus, pour elle, de l’utiliser comme un canal de diffusion qui lui resterait extérieur, mais d’adopter une stratégie qui en intègre le mode de fonctionnement, la logique propre. Elle se dote des outils et des normes¹⁰² qui lui permettent d’être partie prenante de cette évolution, mais de façon plus importante, en adopte les usages pour déterminer sa propre stratégie de visibilité.

La première version de Data.bnf.fr, présentée comme un prototype, ne propose qu’un nombre limité d’œuvres et d’auteurs, en se fondant sur l’existence de notices d’autorité de qualité dans le catalogue. Elle met l’accent sur la création de liens vers les catalogues, Gallica, mais aussi d’autres ressources comme les expositions virtuelles ou les références des sites web archivés, ce qui contribue à faire converger des activités de la bibliothèque qui étaient fortement silotées vers un objectif commun : la visibilité des ressources sur le web. En ce sens, Data.bnf.fr joue, comme SPAR, un rôle de réunification des collections par le biais du numérique. En créant du lien entre différentes ressources, numériques ou non, il fait émerger une vision d’ensemble de ces collections. Et cette vision d’ensemble, parce qu’elle valorise sur le web les données de qualité issues du catalogue, met en relief l’importance du catalogue et donc des activités de traitement et de gestion. À travers une approche orientée données, Data.bnf.fr devient ainsi un trait d’union entre le savoir-faire bibliothéconomique traditionnel et l’univers du web.

102 À ce sujet, voir le rapport du Library Linked Data Incubator Group du W3C, publié en 2011, en ligne : <https://www.w3.org/2005/Incubator/lld/>. Voir également **Emmanuelle Bermès, avec Antoine Isaac et Gautier Poupeau, *Le web sémantique en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2013.

II. Vers l'idée de patrimoine numérique (2012-2014)

En 2011, armée d'un outillage de gestion qui va de la sélection à la valorisation en ligne, en passant par le catalogage et la conservation, la BnF est donc prête à embrasser l'idée de collection numérique. À la fin de l'année est lancée une initiative transverse, associant l'encadrement de la bibliothèque et les experts impliqués dans les projets opérationnels : des groupes de travail, puis un séminaire organisé en 2012 se penchent sur la question du numérique au prisme des grandes missions de la bibliothèque – collecter, conserver, communiquer, valoriser – mais aussi des métiers et des compétences. Le résultat de cette démarche collaborative est un document synthétique intitulé *Numérique horizon 2015*, présenté lors d'une séance d'information au personnel en mars 2013¹⁰³.

Qu'il s'agisse des thématiques qui ont été identifiées pour conduire cette réflexion (politique documentaire, collecte, traitement des collections, conservation, communication, médiation, valorisation) ou des orientations qui en découlent, ce document frappe par l'emploi d'un vocabulaire qui ancre fortement le numérique dans les activités du métier de bibliothécaire et des missions de la bibliothèque, appliquées au numérique. Cette entreprise de légitimation, très englobante, débouche sur l'apparition du terme de « collections numériques », associé à la problématique du signalement et de l'accès, mais aussi de celui de « patrimoine numérique » lorsqu'il est question de conservation. Ainsi, dans les années 2012-2014, l'effort entamé auparavant pour ramener le numérique dans le champ d'une gestion courante de bibliothèque se poursuit autour de différentes thématiques, dont nous retenons ici celles qui témoignent le plus significativement de l'émergence d'un nouveau patrimoine : la désignation de ces nouvelles collections et des actions à entreprendre les concernant, l'évolution des métiers et des compétences, et le développement de la stratégie de présence de la bibliothèque sur le web.

A. Définir de nouveaux objets patrimoniaux

À travers l'initiative *Numérique Horizon 2015*, les collections numériques apparaissent comme multiformes, embrassant à la fois le produit de la numérisation et les documents nati-

103 *Numérique horizon 2015 : synthèse des orientations*. Les midis de l'information. Conférence du lundi 18 mars 2013. BnF, collections audiovisuelles : NUMAV-636013. Voir aussi : *Numérique horizon 2015 : synthèse des orientations*, archives numériques de la BnF : BNF-ADM-2014-101054.

vement numériques, et intégrant toutes les filières, qu'il s'agisse de dépôt légal, d'acquisitions ou de reproduction à des fins de conservation.

Dans le champ de la collecte, la problématique qui émerge est celle de la reprise en main de la sélection, qui fait écho à l'importance de restaurer une politique documentaire, comme nous l'avons vu plus haut. S'il est question d'exhaustivité ou de complétude des collections, la préoccupation porte sur la collecte ou l'acquisition d'objets spécifiques ou unitaires : les titres de presse quotidienne régionale en accès payant qui viendront s'ajouter à la collecte des archives du web, les ebooks qui seront désormais acquis titre à titre par la direction des collections et entreront par dépôt légal, les « varia » qui englobent des objets numériques de toute nature accueillis notamment par les départements spécialisés (archives d'écrivains ou de photographes, fonds de vidéos déposés par des théâtres, mais aussi études de marché et littérature grise). Après avoir traité la masse, on revient en fait plus finement sur ce qui a échappé à la première approche. Dans le domaine de la numérisation, les mêmes années sont marquées par une montée en puissance de la numérisation des collections spécialisées : cartes et plans, estampes et photographies, mais aussi maquettes des décors d'opéra, marionnettes du département des arts du spectacle, collections de monnaies antiques¹⁰⁴...

Ces différentes initiatives, toutes lancées à cette période, ont le point commun de montrer un visage du numérique en forte continuité avec les collections de la BnF dans ce qu'elles ont de plus spécifique. Le don de l'écrivain Pierre Guyotat, qui porte sur l'ensemble de ses archives numériques et sa correspondance par mail et SMS, offre une fenêtre inédite sur un nouveau monde : il permet de se projeter dans ce que représenteraient, demain, la collecte et la conservation d'archives qui seraient l'équivalent numérique des cahiers de brouillons, manuscrits littéraires et lettres d'écrivains que le département des manuscrits conserve en nombre. La collecte de la presse quotidienne régionale, si elle utilise les robots de l'archivage du web, est intégralement cataloguée et bulletinée par les mêmes équipes qui ouvrent chaque matin au service de gestion des périodiques les plis contenant les fascicules papier. Un projet de numérisation des globes en 3D, réalisé grâce au mécénat d'une entreprise japonaise, Dai Nippon Printing, permet de manipuler ces objets dans Gallica presque comme si on les tenait entre ses mains. La chaîne de dépôt légal des livres numériques est développée dans un esprit de mutualisation, notamment sur le plan du signalement, avec le dépôt des livres imprimés.

104 Le volume des collections spécialisées dans Gallica est passé de 100.000 documents numérisés en 2008 à 1,1 million en janvier 2014, d'après Arnaud Beaufort et Florent Palluault, *Évaluation de la politique publique de numérisation des ressources culturelles. Phase 1 : diagnostic*. Rapport du groupe de travail, secteur livre et lecture. Janvier 2014. Archives numériques de la BnF : BnF-ADM-2014-007363, p.26.

À l'heure où l'on commence à parler de patrimoine numérique, désigner de nouveaux objets patrimoniaux interroge donc très directement leur relation aux collections plus traditionnelles, qu'il s'agisse de collecte, de traitement documentaire ou d'accès. La nouvelle *Charte de la conservation*¹⁰⁵, élaborée en 2014-2015, intègre la question de la préservation numérique sans la dissocier de la conservation physique. La BnF explore les frontières des collections numériques à la lumière de ce qu'est la bibliothèque : encore une façon de rétablir l'unité des collections.

B. Faire évoluer métiers et compétences

Dès lors qu'on admet l'existence d'une nouvelle collection sous forme numérique, dont la continuité avec l'ensemble des collections de la BnF est avérée, s'impose la question du métier de bibliothécaire en regard des tâches de gestion de cette collection. Quels sont les changements et les axes de continuité ? Peut-on gérer ces nouvelles collections numériques avec les mêmes savoir-faire et compétences que les collections traditionnelles ? Les travaux autour de SPAR, notamment sur la gestion des risques¹⁰⁶ et la gestion des droits d'accès¹⁰⁷, font émerger le besoin d'une organisation spécifique qui permette de prendre et d'appliquer des décisions qui relèvent du rôle des professionnels des bibliothèques : par exemple, comment déterminer quels documents doivent être préservés en priorité en cas de sinistre ? Peut-on encoder dans un système expert les règles qui déterminent à quelle date un ouvrage tombe dans le domaine public, afin qu'elles s'appliquent automatiquement ? Il s'agit donc d'identifier, voire de former des experts de ces questions et de mettre en place l'organisation afférente.

De 2008 à 2014, l'observatoire ORH-ION¹⁰⁸ poursuit ses travaux pour œuvrer en faveur de cet objectif. On peut compter au nombre de ses réalisations une comparaison entre le circuit de numérisation et celui des archives de l'Internet, l'émergence de nouvelles filières comme celle des acquisitions et dons numériques (qui permet l'acquisition à l'unité, le catalogage, la préservation et l'accès d'objets numériques qui échappent aux circuits des grands éditeurs de revues et livres électroniques), une réflexion sur SPAR et la gestion de la collec-

105 Archives numériques de la BnF : BnF-ADM-2014-087745.

106 Cf. **Emmanuelle Bermès, « Risk management: methodological principles / Gérer les risques : principes méthodologiques ». Dans *International Preservation News* (IFLA-PAC), 2007, n°41, pp. 4-8.

107 Françoise Bourdon, *Implantation organisationnelle du numérique : Gestion des droits : réunions de présentation de SOLON-ORH : 8 et 10 décembre 2008*. Archives numériques de la BnF : BnF-ADM-2008-069892.

108 Cf. **Emmanuelle Bermès et Louise Fauduet, « The Human Face of Digital Preservation... » *op. cit.*

tion numérique, une analyse du profil émergent de magasinier numérique¹⁰⁹. En 2013, lors de la présentation du document stratégique *Numérique Horizon 2015*, le rôle informel de l'observatoire est évoqué lorsque vient la question de la formation ; cependant, il n'est pas cité en tant que tel dans le dispositif décrit dans le texte de synthèse, ce qui confirme sa nature relativement souterraine. L'officialisation de la prise en main de la problématique de l'impact du numérique sur les « identités professionnelles » au sein de la stratégie de l'établissement démontre une évolution de la culture de la bibliothèque, qui commence à asseoir l'idée et les principes d'une gestion des collections numériques.

C. Construire la présence en ligne de la bibliothèque

Face à la masse des collections numériques désormais disponibles, l'enjeu de la sélection documentaire par les bibliothécaires se manifeste également à travers de nouvelles initiatives de médiation, celle-ci s'effectuant a posteriori. Plusieurs programmes sont annoncés à l'occasion de *Numérique Horizon 2015*, visant à éditorialiser des portions de Gallica : les « corpus » qui correspondent à des descriptions de fonds, les « essentiels » qui sont des ensembles de textes majeurs, et la création de points d'entrée géographiques alors intitulés « Gallica territoires ». Ces différentes initiatives, menées de façon transverse, voient le jour dans la bibliothèque numérique en 2013 et 2014, avec l'ajout d'une nouvelle entrée « collections ». Le choix de ce terme n'est pas un hasard : il reflète l'implication de la direction des collections dans ce nouveau dispositif de médiation, et ancre la valorisation des collections numérisées dans les activités de gestion par les bibliothécaires qui ajoutent à leur légitimité. Les parcours guidés dans les archives de l'Internet sont également développés dans cet esprit¹¹⁰.

La valorisation des collections numériques prend aussi d'autres formes, avec le développement d'une très active stratégie de présence de Gallica sur les réseaux sociaux, en particulier Facebook et Twitter, qui vise à démocratiser l'usage des collections numériques. La logique de cette démarche de valorisation se rapproche des codes et des usages du web : en 2010, le partenariat signé avec Wikimedia France¹¹¹ soulève ainsi la question des licences de

109 BnF, *Rapport d'activité 2010*. En ligne : https://multimedia-ext.bnf.fr/pdf/rapport_2010.pdf

110 La liste des parcours guidés dans les archives de l'Internet avec leur date de mise en ligne est disponible sur le site de la BnF : <https://www.bnf.fr/fr/centre-d-aide/parcours-guides-dans-les-archives-de-linternet>. Dix parcours sont ainsi créés entre 2008 et 2014.

111 Il portait sur le versement d'ouvrages dans Wikisource pour en permettre la correction collaborative de l'OCR et sur l'échange de données d'autorités. Si les résultats furent quantitativement décevants, ce par-

réutilisation des documents de Gallica, et la problématique de l'exploitation et de la valorisation des contenus de la BnF sur une plateforme externe.

Enfin, dans la continuité du projet Data.bnf.fr, émerge l'idée d'une médiation « automatique » par le biais de l'ouverture et de la réutilisation des données. La BnF n'est pas la seule institution à se poser à cette période la question de la convergence de ressources hétérogènes autour d'entités clefs que sont les personnes, les œuvres, ou encore les événements, les concepts, les genres... Le Centre Pompidou virtuel, construit suivant les mêmes principes et en utilisant les mêmes technologies, va plus loin car au lieu de se penser comme un pivot chargé de renvoyer l'internaute vers les ressources hétérogènes de l'institution, il tente de centraliser les données pour proposer une nouvelle expérience de navigation, le « parcours de sens »¹¹². Institution multiple dans ses missions et son organisation, le Centre Pompidou présente l'originalité d'agrèger différentes composantes des métiers du patrimoine : un musée, deux bibliothèques dont l'une détient de riches fonds d'archives, et une importante archive audiovisuelle liée à la programmation événementielle. Ainsi, à l'échelle d'une seule institution, la question de la convergence des ressources numériques vers un seul site homogène pose des problèmes similaires à ceux que rencontre, depuis sa création, le portail Europeana qui vise à rassembler la culture européenne sous toutes ses formes : bibliothèques, archives, musées, audiovisuel.

Le web sémantique offre la promesse de préserver les spécificités inhérentes aux différents métiers du patrimoine, telles qu'elles s'expriment dans les données, en autorisant la création de liens entre des descriptions hétérogènes, pour former un espace global d'information autour des personnes, lieux, dates, concepts et œuvres communs à ces différentes cultures professionnelles¹¹³. C'est encore de gestion de collections qu'il s'agit : pour la BnF comme pour le Centre Pompidou, l'enjeu est d'explorer ce que le modèle d'interopérabilité du web, avec l'hypertexte, l'ensemble des normes qui le constitue et

tenariat marqué une étape importante dans l'ouverture de la BnF aux communautés en ligne. Cf. « 1400 ouvrages de Gallica seront versés dans Wikisource. » *Blog Gallica*, 9 avril 2010. En ligne : <http://blog.bnf.fr/gallica/index.php/2010/04/09/1-400-ouvrages-de-gallica-seront-verses-dans-wikisource/>.

112 Cf. **Emmanuelle Bermès, « Des parcours de sens dans le Centre Pompidou virtuel ». Dans *Bulletin des bibliothèques de France*, 2013, n° 5, p. 52-57. En ligne : <http://bbf.enssib.fr/consulter/bbf-2013-05-0052-013>, et **« Following the user's flow in the digital Pompidou », dans Franck Cervone et Lars Svensson (dir.), *Linked Data and User Interaction*, IFLA WLIC 2013 : IFLA World Library and Information Congress, 73rd IFLA General Conference and Assembly. Satellite meeting : Linked Data and User Interaction, Singapour, 2013. Berlin/Boston : De Gruyter/Saur, 2015.

113 Cf. **Emmanuelle Bermès, « Convergence et interopérabilité : l'apport du Web de données ». IFLA WLIC 2011 : IFLA World Library and Information Congress, 77th IFLA General Conference and Assembly. Puerto Rico, 2011. En ligne : <https://www.ifla.org/past-wlic/2011/149-bermes-en.pdf>.

l'expérience de navigation qu'il propose aux internautes, peut apporter à des pratiques documentaires séculaires qu'il oblige à se réinventer.

La double problématique de la convergence institutionnelle et de la réutilisation des données de la BnF dans le contexte du web et des réseaux sociaux se traduit finalement par un questionnement sur la valeur des données, dans le contexte d'un environnement fortement marqué par la culture de l'ouverture et de la gratuité¹¹⁴. Ce qu'on a appelé le « web 2.0 » dans les années 2000, c'est-à-dire le développement d'un outillage simple et gratuit facilitant considérablement la publication mais aussi la réutilisation de contenus sur le web, confronte les institutions culturelles à une question politique qui n'est pas encore pleinement résolue aujourd'hui : quelle est la valeur de la libre circulation des données, et justifie-t-elle de renoncer aux revenus générés par les services auparavant payants ? La BnF a déjà lancé son projet Data.bnf.fr depuis quelques mois lorsque la plateforme Data.gouv.fr voit le jour et qu'avec elle, une politique interministérielle de l'*open data* se développe. Au niveau européen, dans les années 2012-2013, Europeana, héritière directe de la vision de Jean-Noël Jeanneney d'une bibliothèque numérique européenne réunissant face à Google tous les savoirs numérisés dûment organisés en collection de bibliothèque, prône également une forte politique d'ouverture des données. Au-delà du portail Europeana.eu, jugé décevant en raison des limites du modèle OAI-PMH¹¹⁵, la fondation européenne cherche à développer un modèle puissant de réutilisation des données et milite en faveur de la licence CC0¹¹⁶, soit un abandon de toute forme de droits sur les données, y compris l'attribution.

Au terme d'un long débat, la BnF annonce le 1^{er} janvier 2014 le passage sous licence publique de l'État, ou licence Etalab¹¹⁷, de toutes ses métadonnées, et entre ainsi officiellement dans l'*open data*¹¹⁸. L'année 2014 est également celle où la BnF élabore une nouvelle version de sa carte heuristique du *Schéma numérique*, réaffirmant l'expression de sa stratégie dans ce domaine comme un accomplissement naturel de ses missions de gestion des collections : collecter, signaler, conserver, communiquer, diffuser. Le nouveau contrat de perfor-

114 Jean-Noël Jeanneney, *Histoire des médias*. Paris : Seuil, 2011, p. 380.

115 Cf. première partie, chapitre II, B.

116 <https://creativecommons.org/publicdomain/zero/1.0/deed.fr> : licence Creative Commons qui permet de déclarer un document comme appartenant au domaine public, et donc abandonner tous droits de propriété intellectuelle le concernant.

117 <https://www.etalab.gouv.fr/licence-ouverte-open-licence>.

118 Gildas Illien, *Des catalogues de bibliothèque dans le Linked Open Data : servez-vous, la BnF ouvre ses métadonnées*. Cycle enjeux du web 3.0 dans le secteur culturel. Paris : Ministère de la Culture et Inria, 27 janvier 2014. En ligne : http://www.dailymotion.com/video/x1bqoku_servez-vous-la-bnf-ouvre-ses-metadonnees-gildas-illien_news.

mance, conclu avec la tutelle pour la période 2014-2016, s'ouvre sur ces mots : « (...) *la BnF confirmera le rôle moteur qu'elle entend jouer pour intégrer le numérique dans le patrimoine national* ».

III. Collections, corpus, données : la bibliothèque face aux nouveaux usages de la recherche (2014-2019)

En 2014, les collections numériques, désormais reconnues comme telles, et constituées aussi bien d'objets nativement numériques que de reproductions numériques de collections analogiques, sont devenues massives. Le rapport d'activité publié en mars 2015 indique que la collecte large pour le dépôt légal de l'Internet a été la plus importante jamais réalisée, avec un million de nouveaux domaines¹¹⁹. Le plan de sauvegarde des collections audiovisuelles a porté sur près de 100.000 documents vidéo, sonores et cédéroms. SPAR héberge 3,3 millions de documents pour 1 pétaoctet de données et Gallica a dépassé les 3 millions de documents et 15 millions de visiteurs. La BnF est donc bien entrée dans cette « révolution la plus importante depuis Gutenberg », en adaptant toutes ses missions à l'environnement numérique.

Cependant, de même que l'imprimerie naissante a connu avec les incunables une période où les livres sortis des premières presses à caractères mobiles imitaient la mise en page du manuscrit sans exploiter pleinement les potentialités de ce nouveau mode de diffusion, cette première ère numérique s'est attachée en priorité à démontrer et garantir une continuité avec l'ancien monde. La numérisation s'est appliquée à reproduire le plus fidèlement possible les livres originaux, et la bibliothèque numérique Gallica s'est dotée de toutes les fonctionnalités indispensables pour assurer la consultation de chaque document avec tout le confort nécessaire pour ne pas avoir besoin de recourir au document physique. Ce qui va jusqu'à autoriser la lecture de droite à gauche pour les manuscrits en écritures non latines ou à zoomer en haute résolution sur les sceaux avec bien plus de précision qu'on n'aurait pu le faire avec une loupe ! Du côté des archives du web, une « machine à remonter le temps »¹²⁰ rejoue les pages web à l'identique, autant que faire se peut, de leur apparence lorsqu'elles ont été collectées.

À travers ces outils, le numérique démultiplie les possibilités d'accès en autorisant la consultation simultanée sur plusieurs postes par plusieurs lecteurs différents, et à distance lorsque les conditions juridiques le permettent. Mais à ce stade, il ne révolutionne pas –

119 BnF, *Rapport d'activité 2014*. En ligne : https://multimedia-ext.bnf.fr/pdf/rapport_2014.pdf, p. 11.

120 La Wayback Machine, développée par Internet Archive et partagée avec la communauté en *open source* sous le nom d'Open Wayback.

encore – le rapport individuel du lecteur au document, l’usage de la bibliothèque. Celle-ci reste un lieu, physique ou virtuel, où l’usager se rend pour y consulter un ou plusieurs documents.

Cependant, au début des années 2010, une nouvelle évolution se produit : ces énormes masses de données numériques, désormais bien identifiées par les chercheurs, font l’objet d’un intérêt de nature différente dans le contexte d’une évolution de la recherche en sciences humaines : le mouvement qui prend en 2004 le nom de « *digital humanities* », traduit en français par « humanités numériques » en 2010¹²¹. Mouvement qui soulève bien des débats et rassemble une grande diversité de démarches et d’expériences différentes, mais qui, lorsqu’il s’intéresse aux collections, contribue à leur unicité et à la convergence amorcée sur le plan de la conservation et de la gestion, en les traitant comme des données. Ce nouveau paradigme impose, à nouveau, à la bibliothèque un questionnement sur ses missions et une évolution de ses pratiques.

A. Les corpus : une évolution structurelle de la recherche

Pour les chercheurs comme pour la bibliothèque, l’irruption du web a provoqué un bouleversement profond : si depuis les années 1950 l’informatique pouvait servir d’outil de calcul intervenant à un stade précis de la recherche¹²², à l’heure du numérique l’ordinateur est devenu un compagnon de toutes ses étapes, de la collecte des sources à la valorisation des résultats de la recherche, en passant par la recherche bibliographique et l’analyse elle-même. Tous ces aspects ont un impact sur les bibliothèques, parce qu’ils modifient en profondeur le paysage de la production documentaire scientifique : le rôle des grands éditeurs de revues scientifiques numériques dans l’évaluation de la recherche, le mouvement de l’*open access*, le besoin de formation aux outils bibliographiques numériques comme Zotero, la problématique de la préservation des données de la recherche sont autant de sujets qui entrent dans le périmètre des

121 L’ouvrage collectif *A companion to digital humanities* (dirigé par Susan Schreibman, Ray Siemens et John Unsworth) paraît pour la première fois en 2004 ; en France c’est le THATcamp de 2010 qui débouche, en mars 2011, sur le *Manifeste des humanités numériques*. En ligne : <https://tcp.hypotheses.org/318>.

122 Sur les aspects historiques, voir par exemple Pierre Mounier, *Les humanités numériques : une histoire critique*. Paris : Éditions de la Maison des sciences de l’homme, 2018.

nouveaux « services aux chercheurs » développés par les bibliothèques¹²³. Cependant, du point de vue d'une bibliothèque patrimoniale comme la BnF, la question la plus prégnante est celle du changement de la relation des chercheurs à leurs sources, au matériau qu'ils étudient, notamment dans le domaine des sciences humaines et sociales.

En 2006, l'Agence Nationale de la Recherche (ANR) âgée d'un an seulement lance son appel à projets « Corpus » qui vise « *au financement de projets de recherche qui nécessitent la mise en place, la valorisation, l'enrichissement, la diffusion de corpus, ainsi que l'amélioration des outils et des procédures d'analyse* »¹²⁴. La recherche française, désormais largement financée par projets, va donc prendre une part décisive dans la création plus ou moins massive de corpus de sources sous forme numérique. En 2010, d'autres structures de financement liées à l'ANR se créent dans le cadre des investissements d'avenir : les Labex (laboratoires d'excellence) et les Equipex (équipements d'excellence), structures dans lesquelles les institutions patrimoniales, dont la BnF, sont parties prenantes pour apporter à la fois leur expertise et un accès privilégié aux collections. De très grandes infrastructures de recherche, au premier chef la TGIR Huma-Num créée en 2013 par la fusion de deux infrastructures existantes, Adonis et Corpus-IR, ont vocation à soutenir les équipes de recherche en apportant compétences et moyens techniques sous la forme de services. Enfin, en 2014, c'est au tour du réseau des bibliothèques universitaires et de recherche (dont la BnF) de se restructurer dans l'optique des services et de la coopération avec la recherche, avec la préfiguration du dispositif Collex (Collections d'excellence) qui se concrétisera en 2017 par la création d'un GIS (Groupement d'intérêt scientifique). Toutes ces structures reflètent une évolution profonde de la recherche en tant qu'institution, vers un mode qui favorise la mise en réseau, la socialisation des personnels de recherche, la créativité, l'expérimentation et d'une façon générale, la transversalité et l'interdisciplinarité. Elles reflètent un bouleversement des pratiques du même ordre que celui qu'a connu la bibliothèque, avec quelques années de décalage : des tensions et une évolution du travail fortement marquées par la culture numérique qui imprègne de plus en plus la société.

123 Voir les actes de la journée d'étude du congrès ADBU 2017 : *Les bibliothécaires, acteurs de la recherche*. En ligne : <http://adbu.fr/retour-sur-la-journee-detude-du-congres-adbu2017-les-bibliothecaires-acteurs-de-la-recherche/>.

124 Pierre Mounier, *op. cit.* p. 159.

Associée à des projets ANR comme Doremus¹²⁵, aux Labex Les passés dans le présent¹²⁶, Patrima¹²⁷ ou encore Obvil¹²⁸, à l'Equipex Biblissima¹²⁹, la BnF développe une activité importante à la frontière entre pratique professionnelle et recherche, fortement liée aux corpus numériques. Qu'ils conduisent à un travail de modélisation et création de métadonnées (Doremus, Biblissima), à des programmes de numérisation de masse (Obvil), à la mise à disposition de données brutes et réutilisables (Biblissima, Obvil), ces programmes présentent le point commun de poser la question de l'exploitation par les chercheurs de collections numériques massives, existantes ou à créer *ad hoc*. Le point de contact entre bibliothèque et chercheurs, entre collections et corpus, se situe dans la co-construction de nouveaux ensembles de données, comme dans le cas du projet « Le devenir du patrimoine numérisé en ligne : l'exemple de la Grande Guerre »¹³⁰, où une collecte d'archives du web fut réalisée spécifiquement pour permettre l'analyse des réseaux amateurs et institutionnels sur le web. Lorsqu'en 2014, le Labex Obvil émet la demande d'obtenir une copie de Gallica – alors riche de 3 millions de documents – pour effectuer des recherches relevant du « big data », il devient clair que la bibliothèque doit s'organiser pour répondre de manière fluide à ce type de demande. C'est ainsi qu'est imaginé le projet DSR-Corpus, inscrit au plan quadriennal de la recherche de l'établissement pour la période 2016-2019, qui vise à préfigurer de nouveaux services pour les chercheurs autour des collections numériques¹³¹.

125 <https://www.doremus.org/>.

126 <http://passes-present.eu/>.

127 Ce Labex a débouché à sa clôture sur la création de la Fondation des Sciences du Patrimoine : <http://www.sciences-patrimoine.org/>.

128 <http://obvil.sorbonne-universite.site/>.

129 <https://projet.biblissima.fr/>.

130 Conduit de 2013 à 2016 dans le cadre du Labex Les passés dans le présent et porté par la BnF, Télécom ParisTech et la Bibliothèque de documentation internationale contemporaine (BDIC). Cf. Valérie Beau-douin, Philippe Chevallier et Lionel Maurel (dir), *Le web français de la Grande Guerre : réseaux amateurs et institutionnels*. Nanterre : Presses universitaires de Paris Nanterre, 2018.

131 Cf. *Emmanuelle Bermès, « Text, data and link-mining in digital libraries: looking for the heritage gold ». Présentation à l'occasion du Satellite Meeting IFLA *Digital Humanities – Opportunities and Risks: Connecting Libraries and Research*, Berlin, 2017. En ligne : <https://hal.inria.fr/hal-01643293/document>.

B. Les données : de nouvelles formes d'interaction avec les publics

Les premières expérimentations avec les chercheurs autour des données numériques, dont certaines remontent aux origines du projet de numérisation¹³², font apparaître une grande diversité de situations. Alors que dans certaines disciplines, les méthodes scientifiques et l'absence de compétences informatiques au sein des équipes de recherche impliquent un recours aux ressources et outils proposés par la BnF¹³³, d'autres chercheurs ou laboratoires déjà fortement outillés – notamment dans des domaines comme la sociologie, la linguistique ou la recherche en informatique – souhaitent accéder à des données brutes afin de les exploiter dans leurs propres environnements. Certains aspirent à récupérer les données dans des formats simples comme CSV ou JSON¹³⁴, alors que d'autres préfèrent accéder aux formats sources afin de bénéficier des données les plus qualitatives. Certains sont en attente de conseils, d'accompagnement, d'expertise alors que d'autres souhaitent télécharger les données en toute autonomie sans avoir l'obligation d'interagir avec les bibliothécaires.

La diversité des demandes et des situations conduit la BnF, dans les années 2016-2017, à deux constats : d'une part, il est important de multiplier les canaux d'accès aux données sous différentes formes pour créer des opportunités de réutilisation, d'autre part la bibliothèque doit être prête à s'engager dans un dialogue « sur mesure » avec les équipes de recherche qui travaillent sur le numérique, pour les accompagner sur différents aspects : juridique (qu'ai-je le droit de faire avec les données ?), technique (comment comprendre les formats ?), scientifique (quels contenus puis-je m'attendre à trouver, comment les interpréter ?).

Le projet DSR-Corpus se construit ainsi par itérations successives, en expérimentant à taille réelle, avec des équipes de recherche partenaires, les nouveaux services envisagés. Deux autres initiatives méritent d'être mentionnées. Pour la première fois en 2016, un hackathon de 24 heures est organisé : la BnF ouvre ses portes à des équipes de développeurs, chercheurs, étudiants, designers et autres passionnés pour imaginer et prototyper de nouveaux projets au-

132 On a mentionné en première partie les projets IDA et Optima, on peut aussi parler de Stitch, projet conduit avec l'Université Libre d'Amsterdam, qui fut la première expérimentation de la BnF dans le domaine du web sémantique : <https://c.bnf.fr/DYE>.

133 Cf. Sophie Gebeil, « Quand l'historien rencontre les archives du web ». Dans *Web Corpora*, 10 novembre 2017. En ligne : <https://webcorpora.hypotheses.org/380> et « Quand l'historien rencontre les archives du Web », *Revue de la BnF*, 2016 n°2, p. 185-191.

134 CSV (Comma-Separated value) et JSON (JavaScript Object Notation) sont deux formats courants d'échange de données.

tour de ses collections numériques. Cette initiative, reconduite en 2017 et 2018, fait apparaître l'intérêt d'un accompagnement de proximité autour de l'usage des données de Gallica, de Data.bnf.fr ou autres, afin de donner aux usagers toutes les clefs dont ils ont besoin pour les réutiliser à pleine puissance. En 2017, lors de la deuxième édition du hackathon, un site dédié à la documentation de ces accès techniques est créé : BnF API et jeux de données¹³⁵. Accès aux images de Gallica via l'API « IIF »¹³⁶, téléchargement de jeux de données en RDF¹³⁷ issus de Data.bnf.fr, interrogation du catalogue en utilisant le protocole SRU¹³⁸, récupération des listes de « graines » des collectes du web : ces différentes modalités sont désormais réunies en un seul point d'entrée, qui valorise et documente un usage des collections numériques et des données de la BnF déconnecté des interfaces de lecture classiques. En outre, sur le site API et données, sont également redistribuées les données enrichies issues de projets de recherche : par exemple, les 135.000 ouvrages de Gallica transformés en TEI par le Labex Obvil, ou le corpus de presse numérisée enrichie dans le cadre du programme européen European Newspapers.

C. Convergence et questions épistémologiques

Dans le cadre du projet DSR-Corpus, à travers le site API et données ou lors des hackathons, les collections numériques de la BnF sont exploitées suivant un nouveau paradigme, caractérisé dans le monde anglo-saxon par le terme « *collections as data* »¹³⁹ que l'on pourrait traduire en français par la périphrase hélas peu satisfaisante « des collections exploitées comme des données ». Cette approche s'insère dans un contexte plus global d'importance croissante de la notion de données qu'on a pu qualifier de « dataification du monde » : essor de la datascience, du big data, du data journalisme... sont des tendances qui marquent la so-

135 API : Application Programming Interface. Cf. <http://api.bnf.fr>.

136 Cf. **Emmanuelle Bermès et Matthieu Bonicel, « Biblissima and IIF ». Présentation pour IFLA WLIC 2016 : IFLA World Library and Information Congress, 82nd IFLA General Conference and Assembly. Columbus (Ohio), 2016. Voir aussi : Emmanuelle Bermès et Stéphane Pillorget, « Gallica et IIF : vers un nouveau modèle d'interopérabilité ». Présentation pour la journée Biblissima et IIF *Innover pour redécouvrir le patrimoine écrit*, 15 mars 2018 (Campus Condorcet). En ligne : <https://projet.biblissima.fr/fr/gallica-iif-vers-nouveau-modele-interoperabilite>.

137 Resource Description Framework.

138 Search/Retrieve via URL.

139 Thomas Padilla, Laurie Allen, Hannah Frost, Sarah Potvin, Elizabeth Russey Roke, Stewart Varner, *Final Report – Always Already Computational: Collections as Data*. En ligne : <https://zenodo.org/record/3152935>.

ciété de manière globale¹⁴⁰. Ces « collections-données » sont définies par leur usage : le fait d'exister sous forme numérique les rend propres à une exploitation informatique qui inclut la fouille de texte, la visualisation de données, la cartographie, l'analyse d'images, de sons et de réseaux¹⁴¹. Cette définition peut donc s'exercer sur toute forme de collections numériques, qu'elles soient issues de la reproduction d'objets analogiques ou nativement numériques, et quel que soit leur contenu intellectuel. À la BnF, une telle définition permet d'englober dans le concept de « collections-données » non seulement la numérisation de tous types de documents, les acquisitions électroniques, les collections audiovisuelles et multimédia, les archives de l'Internet et autres documents numériques collectés par dépôt légal, mais également les données des catalogues, qui peuvent devenir elles-mêmes objets de recherches, et d'autres types de données administratives ou techniques, comme les traces d'usage ou « logs » de Gallica¹⁴².

Cependant, ces nouvelles approches font émerger des questions épistémologiques dont la bibliothèque ne peut qu'être partie prenante. L'analyse et l'interprétation des résultats, souvent quantitatifs ou même, dans le cas de la datavisualisation, essentiellement sensoriels, requièrent une compréhension fine de la manière dont le corpus a été constitué¹⁴³. Cette question est d'autant plus prégnante que les règles qui président à la constitution des collections numériques pré-existent à l'élaboration du questionnement scientifique : les chercheurs sont donc confrontés à un corpus élaboré par d'autres personnes et à d'autres fins que celui de leur recherche¹⁴⁴. En octobre 2017, l'historien médiéviste Paul Bertrand concluant la journée « Humanités numériques et données patrimoniales : publics, réseaux, pratiques »¹⁴⁵ appelait les professionnels des institutions patrimoniales à mobiliser l'expertise accumulée en plus de vingt ans de pratique du patrimoine numérique pour élaborer avec les chercheurs une « refonte de la critique documentaire », qui permettrait de mieux appréhender ce nouveau para-

140 Dominique Cardon, *À quoi rêvent les algorithmes ?* Paris : Seuil, 2015.

141 *The Santa Barbara statement on Collections as data*, 2017. En ligne : <https://collectionsasdata.github.io/statement/>.

142 Adrien Nouvellet, Valérie Beaudouin, Florence D'Alché-Buc, Christophe Prieur, François Roueff, *Analyse des traces d'usage de Gallica : Une étude à partir des logs de connexions au site Gallica*. Rapport de recherche. Paris : Télécom ParisTech ; Bibliothèque nationale de France, 2017. En ligne : <https://hal.archives-ouvertes.fr/hal-01709264>.

143 P. Mounier, *op. cit.*, p. 97.

144 Valérie Beaudouin, dans *Le web français de la Grande Guerre... op. cit.*, p. 117-118.

145 *Conférence conclusive : le troisième oeil*. En ligne : <https://www.youtube.com/watch?v=9BZCGS18BR8>.

digne de la donnée et la transversalité inédite entre les disciplines qui en découle, qu'il s'agisse de méthodologie, d'outils ou de corpus. Une telle approche consacre la reconstitution de l'unité des collections numériques, sous l'étendard des usages. Ceux-ci plaident en faveur du développement d'une science de la donnée qui ne soit pas purement technique, mais documentaire. Le numérique s'inscrit dès lors en continuité des collections dans leur ensemble, tant pour les chercheurs face à leurs corpus de sources que pour les bibliothécaires dans leurs tâches de gestion.

Troisième partie :

Le processus de patrimonialisation du numérique en
bibliothèque

En bibliothèque, la notion de patrimoine ou de fonds patrimonial reste encore aujourd'hui fortement corrélée à celle de documents « anciens, rares et précieux », qui figure dans les textes officiels¹⁴⁶. Tout en étant généralement reconnue par la profession comme insuffisante et imprécise, cette définition reflète l'histoire de la constitution des fonds anciens dans les bibliothèques, essentiellement issue des confiscations révolutionnaires, et le développement à partir du début du XIX^e siècle des « réserves » destinées à abriter les ouvrages qui étaient identifiés comme à préserver avec un soin particulier. Répondre à l'un de ces trois critères – l'ancienneté, la rareté ou le prix – suffisait à qualifier les documents comme « patrimoniaux », ce qui a permis d'étendre les « fonds anciens » des bibliothèques de tous statuts en y adjoignant notamment des ressources d'intérêt local (que celui-ci soit géographique ou institutionnel) et des objets autres que des livres (manuscrits, archives, fonds iconographiques ou autres).

Lorsqu'au XX^e siècle les bibliothèques ont développé leurs missions éducatives, culturelles et documentaires, une opposition de fait a émergé entre celles-ci et les missions de conservation des collections patrimoniales¹⁴⁷. Cette scission a eu pour conséquence de renforcer la définition la plus étroite du patrimoine en bibliothèque, au regard des enjeux de traitement spécifique et de compétences *ad hoc* qui exigeaient une gestion à part des fonds concernés¹⁴⁸. On se trouvait dès lors d'un côté avec des collections dont l'exigence de conservation justifiait une gestion adaptée et qui offraient un fort potentiel de valorisation (par exemple par des expositions ou par la numérisation), et de l'autre avec des collections d'usage courant caractérisées par une forte rotation et qui n'étaient pas destinées à être conservées à moyen ou long terme.

Si l'on retient cette définition classique, les bibliothèques présentent donc, par rapport à d'autres institutions patrimoniales comme les musées, la particularité d'avoir la charge de collections qui sont pour partie patrimoniales, pour partie courantes et dispensables. Il faut

146 Décret n°2001-650 du 19 juillet 2001, cité par Raphaëlle Mouren dans *Manuel du patrimoine en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2007, p. 22. Voir aussi Louis Desgraves, *Le patrimoine des bibliothèques : rapport à M. le directeur du livre et de la lecture*. Paris : Ministère de la Culture, 1982. Voir aussi Jean-Paul Oddos (dir.), *Le Patrimoine. Histoire, pratiques et perspectives*. Paris : Éditions du Cercle de la librairie, 1997.

147 Agnès Marcetteau-Paul, « Le Patrimoine, une valeur d'avenir ? ». Dans *Bulletin des bibliothèques de France*, 2004, n° 5, p. 35-38. En ligne : <http://bbf.enssib.fr/consulter/bbf-2004-05-0035-006>.

148 On peut mentionner par exemple le dispositif de mise à disposition de conservateurs d'État dans certaines bibliothèques territoriales, visant à l'origine la gestion de ces fonds particuliers que l'État avait délégué aux collectivités à la suite des confiscations révolutionnaires.

donc supposer qu'un processus de patrimonialisation y est à l'œuvre, comparable à celui du classement des monuments historiques ou du référencement opéré par l'Inventaire national, faisant entrer chaque objet concerné dans « *un nouveau régime d'existence qui exige, explicitement, sa pérennité intégrale* »¹⁴⁹. Ce processus de patrimonialisation, observé depuis plus de deux siècles, a été largement étudié, par différents auteurs, qu'ils soient anthropologues, historiens, sociologues ou chercheurs en sciences de l'information¹⁵⁰. Nous proposons de nous appuyer sur les outils théoriques qu'ils ont développés pour tenter de saisir l'existence, à la BnF, d'un nouveau patrimoine numérique et de ses caractéristiques propres.

Trois pistes s'offrent à nous dans cette analyse : la première, institutionnelle, porte sur la définition de procédures et de règles qui déterminent la nature des objets susceptibles d'appartenir au patrimoine. La deuxième, opérationnelle et professionnelle, repose sur les différentes étapes du processus de patrimonialisation. La troisième enfin nous emmènera sur le terrain des « émotions patrimoniales » qui sous-tendent la relation intime de chacun de nous à un patrimoine dans lequel il se reconnaît.

149 Daniel Fabre, « Le patrimoine porté par l'émotion. » Dans Daniel Fabre (dir.), *Émotions patrimoniales*. Paris : Éditions de la Maison des Sciences de l'Homme, 2013, pp. 58. La plupart des références à D. Fabre dans notre propos font référence à ce texte.

150 On se référera notamment aux travaux de Daniel Fabre, Dominique Poulot, Nathalie Heinich, Jean Davallon et autres auteurs cités ci-dessous au fil de notre argumentation.

I. La patrimonialisation institutionnelle du numérique : de la reconnaissance à la légitimité

À l'heure de définir l'existence et la nature d'un patrimoine numérique à la BnF, la première question qui se pose est celle de sa légitimité. Or, la valorisation classique par la profession des documents « anciens, rares ou précieux » soulève à cet égard des problèmes évidents : malgré les incertitudes et débats qu'a pu poser la question de la barrière chronologique qui permettrait d'identifier de manière certaine les fonds « anciens », les collections numériques ne peuvent entrer dans cette catégorie ; tout au plus sont-elles pour partie des reproductions de documents anciens, ce qui ne suffit pas à leur conférer une nature patrimoniale. Le numérique ayant intrinsèquement la propriété d'être reproductible à l'envi, il est également impossible de le qualifier de « rare ». Enfin, ces objets qui s'inscrivent dans la dynamique de libre circulation de l'information propre à l'environnement du web sont difficiles à valoriser économiquement, et s'il est possible de leur associer une notion de coût (celui qu'a représenté leur mise à disposition au sein des collections), on ne peut en dire autant de la notion de valeur.

Faire entrer les collections numériques dans le périmètre des documents « anciens, rares et précieux » que les bibliothèques considèrent habituellement comme des fonds patrimoniaux nous obligerait donc à étirer considérablement l'acceptation de ces trois notions, jusqu'aux limites de la logique, en défendant l'hypothèse d'une ancienneté, d'une rareté et d'une préciosité entièrement relatives, qui se mesureraient de façon spécifique pour le numérique, sans lien avec l'application de ces critères à d'autres types de collections : une voie qui ne nous semble pas pertinente.

Ainsi, le numérique en bibliothèque ne pouvant prétendre à l'appellation rassurante d'« ancien, rare et précieux » validée dans la profession, par quels outils peut-il accéder au statut de patrimoine ? Dans *Le don du Patrimoine*, Jean Davallon développe l'hypothèse selon laquelle les objets patrimoniaux semblent l'être par nature aux yeux des professionnels, parce qu'ils ont déjà été patrimonialisés d'un point de vue institutionnel, à l'aide de règles et procédures définies¹⁵¹. En premier lieu, il nous faut donc considérer ici les outils d'institutionnalisation qui officialisent le statut du patrimoine numérique indépendamment du détail des objets visés. Mais aussitôt, une autre question émerge : ces dispositifs sont-ils suffisants pour légitimer le numérique en tant que patrimoine ?

151 Jean Davallon, *Le don du patrimoine*. Paris : Hermès science publications, Lavoisier, 2006, p. 127 et ss.

A. Du patrimoine écrit au patrimoine numérique

En 2004, le *Bulletin des Bibliothèques de France* consacre un dossier thématique aux Patrimoines – on notera l’emploi du pluriel : premier indice d’une remise en cause de la définition traditionnelle du patrimoine en bibliothèque ? En effet, si la notion de patrimoine telle que nous l’utilisons aujourd’hui a émergé au moment de la Révolution française, elle a connu une première mutation dans la deuxième moitié du XX^e siècle, puis une véritable « explosion patrimoniale » dans les années 1980 et 1990¹⁵². Dans son éditorial, Anne-Marie Bertrand évoque cette « patrimonialisation » qui pourrait conduire à faire de tout, ou de n’importe quoi, un objet patrimonial, pour poser aussitôt la question du patrimoine écrit, dans son acception classique, et des ressources que sa préservation exige. Quelques lignes qui suffisent à éclairer la situation des bibliothèques : face à cette « explosion patrimoniale » qui dure depuis deux décennies, celles qui détiennent des fonds « anciens, rares et précieux » s’interrogent sur les moyens qui sont mis à leur disposition pour en assurer la garde, en concurrence avec d’autres missions, au service d’un public plus large.

Daniel Fabre qualifie d’« âge du monument » la période qui, jusqu’en 1980, faisait prévaloir une définition du patrimoine déterminée par des autorités, à la fois scientifiques et gouvernementales, et disséminée par le biais de l’école de la République. Le patrimoine est alors constitué de « monuments » c’est-à-dire d’œuvres remarquables par leur esthétique comme leur histoire, qu’il s’agisse d’architecture, de peinture ou de littérature. À partir des années 1980, en revanche, on entre dans l’« âge du patrimoine », où celui-ci est choisi par les communautés qui se reconnaissent en lui et l’élèvent à un rang particulier. Ainsi, dans le même numéro du *Bulletin des Bibliothèques de France*, on trouve sous la plume de Michel Melot l’assertion suivante : « *L’objet patrimonial est simplement l’objet grâce auquel une communauté existe. Elle a besoin de lui pour exister, et il a besoin d’elle pour exister en tant que patrimoine. (...) Il transforme les populations en Peuples et les territoires en Nations.* » Le fait que Michel Melot ne soit pas seulement conservateur des bibliothèques mais qu’il ait travaillé plusieurs années à l’Inventaire peut expliquer en partie qu’il porte sur la notion de patrimoine un regard complètement différent de l’approche traditionnelle de la profession. Cependant, la lecture des autres articles confirme qu’il s’agit d’une tendance nouvelle, qui va de

152 Jacques Le Goff, *Patrimoine et passions identitaires*. Entretiens du patrimoine, Théâtre national de Chail- lot, Paris, 6 au 8 janvier 1997. Paris : Fayard, Éditions du Patrimoine, 1998, p. 10.

manière évidente faciliter grandement la reconnaissance des collections numériques comme patrimoniales.

La *Convention pour la sauvegarde du patrimoine culturel immatériel* adoptée par l'Unesco en octobre 2003¹⁵³ institutionnalise l'évolution d'une définition du patrimoine fondée sur les communautés et non plus sur l'érection de « monuments » par des autorités scientifiques ou politiques. En effet, elle définit explicitement le patrimoine comme ce « *que les communautés, les groupes et, le cas échéant, les individus reconnaissent comme faisant partie de leur patrimoine culturel* ». Cette définition tautologique, qui présente l'intérêt d'admettre l'absence d'universalité de la notion, renvoie de fait aux communautés la charge d'élire leur propre patrimoine.

Or, c'est à l'occasion de la même conférence générale de l'Unesco, en 2003, qu'est adopté un autre texte : la *Charte de la conservation du patrimoine numérique*¹⁵⁴. Si le patrimoine immatériel vise la culture orale et plus largement tout ce qui relève du vivant¹⁵⁵, le patrimoine numérique apparaît clairement comme un héritier direct du patrimoine écrit :

Le patrimoine numérique se compose de ressources uniques dans les domaines de la connaissance et de l'expression humaine, qu'elles soient d'ordre culturel, éducatif, scientifique et administratif ou qu'elles contiennent des informations techniques, juridiques, médicales ou d'autres sortes (...) Les documents numériques comprennent, parmi un large éventail de formats électroniques qui ne cesse de se diversifier, des textes, des bases de données, des images fixes et animées, des documents sonores et graphiques, des logiciels et des pages web.

La charte distingue les ressources « *d'origine numérique* » – qui sont à préserver en priorité – de celles qui sont « *converties sous forme numérique à partir de ressources analogiques existantes.* » Elle précise enfin que « *la législation en matière d'archives et de dépôt légal ou volontaire dans des bibliothèques, archives, musées et autres dépôts publics doit être étendue au patrimoine numérique* ». Il s'agit donc bien de reconnaître l'existence d'un patri-

153 En ligne : <https://ich.unesco.org/fr/convention>.

154 En ligne : https://unesdoc.unesco.org/ark:/48223/pf0000133171_fre.page=89.

155 À savoir : les traditions et expressions orales, y compris la langue ; les arts du spectacle ; les pratiques sociales, rituels et événements festifs ; les connaissances et pratiques concernant la nature et l'univers ; les savoir-faire liés à l'artisanat traditionnel.

moine documentaire sous forme numérique et de recommander sa conservation au moyen de procédures institutionnelles¹⁵⁶.

Trois ans plus tard, en 2006, cette recommandation se concrétise en France par la loi dite « DADVSI » (Droits d'Auteurs et Droits Voisins dans la Société de l'Information) qui étend le dépôt légal aux « *signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique* ». Lorsqu'en 2008, le *Bulletin des Bibliothèques de France* revient sur la question patrimoniale, cette fois sous le titre « *Construire le patrimoine* », il y est question de numérisation, d'archivage numérique, d'archivage de l'Internet, de dématérialisation du patrimoine sonore, de catalogage informatisé des manuscrits... En quatre ans, le numérique a fait sa place dans un « patrimoine » des bibliothèques redevenu unifié.

En 2015, une nouvelle recommandation de l'Unesco intitulée *Recommandation concernant la préservation et l'accessibilité du patrimoine documentaire, y compris numérique*¹⁵⁷ vient confirmer cette perception de continuité entre patrimoine documentaire et patrimoine numérique, faisant de ce dernier, officiellement et légitimement, l'héritier du patrimoine écrit ancien, rare et précieux. Un héritage à deux facettes, puisqu'il s'agit à la fois de conserver le patrimoine documentaire né numérique qui circule sur les réseaux, et d'utiliser les technologies numériques pour diffuser et valoriser le patrimoine plus traditionnel.

B. Le dépôt légal, outil de patrimonialisation institutionnelle

Le Code du Patrimoine fait de la BnF l'un des trois établissements dépositaires du dépôt légal, avec l'Institut national de l'audiovisuel (Ina) et le Centre national du cinéma et de l'image animée (CNC). Mission régaliennne de la Bibliothèque nationale de France depuis l'ordonnance de Montpellier de 1537, le dépôt légal vise à constituer un patrimoine national représentatif de la production éditoriale de son temps. À l'obligation de dépôt pour les éditeurs s'assortit, pour la bibliothèque, celle de signaler ce patrimoine dans la *Bibliographie nationale française* et d'y donner accès, dans ses emprises, aux chercheurs accrédités. À travers le temps, le dépôt légal s'est adapté aux évolutions des formes et supports de la production éditoriale. Deux dates clefs marquent ainsi l'intégration du numérique à ce dispositif : 1992, avec la loi intégrant au dépôt légal les premiers documents électroniques sur support

156 B. Dufrière, *Heritage and digital humanities... op. cit.* p.10.

157 En ligne : <https://unesdoc.unesco.org/ark:/48223/pf0000244675>.

(jeux vidéo, logiciels et bases de données) et 2006, date à partir de laquelle le dépôt légal embrasse désormais le web mais aussi à terme les autres formes de publications dématérialisées (ebooks, plateformes musicales, vidéo à la demande, etc.).

L'objectif du dépôt légal est de constituer un patrimoine national représentatif et autant que faire se peut, exhaustif de la production de la nation, qu'elle soit scientifique, littéraire, artistique, intellectuelle ou commerciale, sous toutes ses formes. Aucun jugement de valeur, qu'il soit moral, esthétique ou social, n'est appliqué pour sélectionner les documents appelés à faire patrimoine, au contraire d'une politique documentaire classique¹⁵⁸. Ainsi, en théorie, le seul fait d'étendre le périmètre du dépôt légal au numérique devrait suffire à légitimer le statut patrimonial de ces nouveaux objets intégrant les collections.

En réalité, les choses ne sont pas aussi simples. Si le dépôt légal tend à l'exhaustivité, il se heurte en réalité à de nombreuses limitations pratiques : réticence des déposants ou ignorance, de leur part, de l'existence du dispositif ; impossibilité technique de collecter, conserver ou signaler certains objets ; délimitation d'un périmètre d'exclusion pour des objets apparentés à des documents mais qui ne sont en réalité pas de nature documentaire¹⁵⁹. Ces limitations ont conduit la bibliothèque à adapter ses pratiques et opérations de traitement : ainsi, certains documents considérés comme étant d'une pertinence moindre pour la recherche, comme par exemple certains manuels scolaires, sont traités « en recueils » – c'est-à-dire qu'ils ne sont pas catalogués à la pièce – tandis que d'autres, comme la production relevant d'un champ qui a été caractérisé comme « autoédition », subissent un catalogage allégé et ne font pas l'objet de veille éditoriale ni de relance en cas de non-dépôt par l'éditeur. Or, derrière ces choix, on décèle bien, si ce n'est un jugement de valeur, du moins une forme de classement qui différencie les objets patrimoniaux en fonction de leur potentiel d'usage en tant que documents de bibliothèque.

L'irruption du numérique dans le champ du dépôt légal a un impact certain sur ces subtiles variations dans les pratiques autour d'un champ patrimonial traditionnellement présenté comme exhaustif et homogène. Tout d'abord, il implique une massification de la collection, y

158 Cf. **Emmanuelle Bermès, « Éthique et numérisation : les enjeux éthiques de la diffusion numérique. » Dans Géraldine Goffeaux Callebaut (dir.), *Éthique et patrimoine culturel, regards croisés*. Paris : L'Harmattan, 2016, p. 219-228.

159 Par exemple des carnets vierges ou des agendas qui sont déposés par des éditeurs de livres s'étant diversifiés dans la papeterie.

compris sous forme papier dans la mesure où il n'a jamais été aussi facile de produire des livres¹⁶⁰, mais aussi dans d'autres domaines comme la musique¹⁶¹.

Cependant, cette massification s'accompagne pour partie de la possibilité d'automatiser certains processus : collecte automatique, dépôt systématique par flux, récupération de métadonnées des éditeurs... Si bien que le numérique peut aussi apparaître comme un moyen de rendre le dépôt légal plus exhaustif qu'il ne l'a été auparavant. Ainsi, la collecte de la presse quotidienne régionale par le biais des outils d'archivage du web a permis, pour les titres concernés, de collecter sous forme de fichiers PDF la totalité des éditions locales, alors qu'auparavant, n'étaient conservées sous forme papier et microfilm que les éditions des plus grandes villes. Des titres comme *Ouest France* ou les *Dernières nouvelles d'Alsace* intègrent donc désormais les collections patrimoniales à raison de plusieurs centaines de variantes locales, constituant une source extrêmement riche pour la recherche à l'avenir, par exemple dans le domaine de la généalogie.

Dans ce cas précis, la légitimité patrimoniale de l'objet numérique – un fichier correspondant à un fascicule de presse – découle sans difficulté de l'évidente continuité avec la collection traditionnelle. On a donc affaire à un simple changement de support qui, s'il modifie les procédures à différents points de la chaîne de traitement, ne pose aucune question de légitimité. Or, ce n'est pas le cas de tous les documents numériques : dans sa thèse de doctorat¹⁶², Benjamin Barbier a ainsi montré qu'avant que les jeux vidéo entrés dans les collections par dépôt légal ne soient reconnus comme « patrimoine vidéoludique », un certain nombre d'opérations ont dû être entreprises, parmi lesquelles une valorisation scientifique visant à démontrer l'intérêt de ces objets comme sources pour les sciences humaines et l'histoire culturelle. On peut néanmoins se demander avec Benjamin Barbier si la question posée ici est celle de leur légitimité en tant que patrimoine ou simplement en tant qu'objet culturel. Dans le deuxième cas, l'ambiguïté découlerait donc de leur statut d'objet de culture de masse et non de leur nature numérique, impliquant une valeur scientifique moindre, au même titre que l'autoédition comme nous l'avons vu plus haut. Ces objets se trouvent alors « en attente » d'une valeur patrimoniale additionnelle qui leur serait apportée plus tard par l'ancienneté, la rareté ou la préciosité, ou par leur reconnaissance en tant que source d'un travail scientifique :

160 Cf. **Emmanuelle Bermès, « Quand le dépôt légal devient numérique : épistémologie d'un nouvel objet patrimonial. » Dans *Quaderni*, *op.cit.*

161 Le volume de publications dématérialisées sur les plateformes payantes de musique en ligne était évalué en 2017 à 180.000 albums par an. Cf. BnF, *Contrat d'objectifs et de performance 2017-2021*.

162 Benjamin Barbier, *op. cit.*

la collection des livres autoédités ou celles des jeux vidéo contiennent peut-être l'œuvre de jeunesse d'un nouveau Marcel Proust ou d'un futur Quentin Tarantino, ce qui est impossible à prédire et oblige l'institution à conserver la totalité des objets en attendant qu'émerge la « trouvaille »¹⁶³. La patrimonialisation institutionnelle au titre du dépôt légal intervient alors à titre préventif, afin de s'assurer que ces objets ne disparaissent pas en totalité dans la période potentiellement très longue qui peut s'écouler avant qu'ils ne commencent à éveiller ce type d'intérêt.

C. L'extension progressive du domaine du patrimoine numérique

Un autre élément à prendre en compte dans notre réflexion est la spécificité de la BnF dans le paysage des bibliothèques, en France comme dans le monde. À la fois bibliothèque de recherche, fréquentée majoritairement par des étudiants et des chercheurs, institution historiquement dépositaire des anciennes collections royales, et instrument d'affirmation de l'État-Nation à travers sa mission de dépôt légal, elle se définit comme « patrimoniale » à plus d'un titre. La logique de scission entre cette mission patrimoniale et les fonctions d'une bibliothèque d'étude telle qu'elle était conçue au XX^e siècle, évoquée plus haut, a conduit à la considérer longtemps comme une bibliothèque « de dernier recours », dont la préoccupation devait se focaliser sur la conservation d'ouvrages dont elle était, pour beaucoup, seule dépositaire. La consultation et la conservation étant, dans le domaine du livre, difficilement compatibles, cette vision conduisait logiquement à réserver l'accès aux salles de lecture de l'ancienne Bibliothèque nationale à un public restreint.

L'idée de « très grande bibliothèque, d'un genre entièrement nouveau » puis, en 1994, la fusion entre l'établissement public Bibliothèque de France et la Bibliothèque nationale mirent *de facto* fin à cette conception de l'institution. Certes, la césure entre bibliothèque de recherche, au Rez-de-jardin du nouveau bâtiment, et bibliothèque d'étude en Haut-de-jardin – césure qui se traduit jusque dans les collections, entre lesquelles il n'existe aucune porosité – avait vocation à prolonger une distinction claire entre les deux missions, conservation et recherche d'un côté, étude ouverte à tous publics de l'autre. Cependant, l'entrée de la BnF dans l'ère du numérique était appelée à bouleverser en profondeur la scission, voire à la renverser totalement. Le « patrimoine » (c'est-à-dire les documents anciens, rares et précieux) désor-

163 L'idée de trouvaille dans le patrimoine écrit est magnifiquement illustrée sous la plume d'Arlette Farge dans *Le goût de l'archive* (Paris : Seuil, 1989). Voir le chapitre suivant pour le rôle de la notion de « trouvaille » dans le processus de patrimonialisation.

mais numérisé et majoritairement dans le domaine public, est soudain accessible à tous, à distance, y compris à des fins d'étude et de recherche et plus seulement de valorisation. Les « chercheurs » n'ont donc plus besoin de fréquenter la bibliothèque physiquement pour accéder à ces trésors¹⁶⁴. En revanche, les fonds nécessaires à l'« étude », plus contemporains, entrés par dépôt légal ou acquis sous forme numérique avec des licences parfois restrictives, ne pouvant être diffusés aussi largement, perpétuent la fonction de lieu de travail et d'accès aux collections de la bibliothèque.

Face à ces différentes conceptions de la bibliothèque et de ses publics, c'est l'unité des collections¹⁶⁵ qui l'emporte et conduit à assimiler l'ensemble des documents à ceux qui sont patrimoniaux par nature. Si les ouvrages mis à disposition du public en Haut-de-jardin peuvent faire l'objet d'un désherbage, ceux des magasins, bien qu'entrés par acquisitions, répondent aux mêmes exigences de préservation que ceux qui sont issus du dépôt légal. Un phénomène similaire a été observé dans le cas des collections numériques : si SPAR offrait la possibilité de définir, via ses différentes filières, des politiques de préservation diversifiées et adaptées en particulier au contexte juridique de l'entrée des documents – les filières de dépôt légal étant distinctes des filières d'acquisition, la numérisation de diffusion étant séparée de la numérisation de conservation – en réalité, le cycle de vie du paquet d'information dans SPAR a été conçu à l'origine pour une conservation absolue et ne prévoyait pas la suppression. Cette fonctionnalité a été ajoutée par la suite pour les besoins du tiers archivage¹⁶⁶.

On assiste donc à une sorte de patrimonialisation par capillarité : les collections qui ne sont pas de nature patrimoniale par essence le deviennent par l'effet de leur finalité commune et de pratiques de gestion uniformisées. Leur seule présence au sein de l'institution leur confère une légitimité patrimoniale.

164 Mélanie Roustan, « La vague numérique et le ressac du réel. » Dans *La recherche dans les institutions patrimoniales : sources matérielles et ressources numériques*. Villeurbanne : Presses de l'ENSSIB, 2016, p. 175.

165 Voir partie 2.

166 Bertrand Caron, Jordan de la Houssaye, Thomas Ledoux et Stéphane Reecht, « Life and Death of an Information Package: Implementing the Lifecycle in a Multi-Purpose Preservation System ». Présenté à iPRES 2017 : 14th International Conference on Digital Preservation, 2017, Kyoto (Japon). En ligne : <https://hal-bnf.archives-ouvertes.fr/hal-01617645/>.

II. Le processus de patrimonialisation numérique en questions

Les auteurs qui ont travaillé sur le processus de patrimonialisation se sont penchés d'une part sur les critères qui pouvaient conduire à déterminer la nature patrimoniale des objets visés, d'autre part sur l'identification des différentes étapes du processus. Concernant les critères, nous avons vu précédemment que l'extension du dépôt légal, l'unité des collections et la revendication d'une approche globale de celles-ci conduisent à une définition très étendue de la nature patrimoniale des collections numériques, liée à l'institution elle-même. Une étude plus approfondie des étapes de la patrimonialisation pourrait donc compléter cette vision et nous fournir davantage d'indices sur le statut patrimonial des collections numériques.

Dans *Le don du patrimoine*, Jean Davallon identifie ainsi six « gestes » constitutifs du processus de patrimonialisation : la découverte de l'objet comme « trouvaille », la certification de son origine, la confirmation de l'existence du monde d'origine de l'objet, la représentation de ce monde par l'objet, la célébration de l'objet par l'exposition, et enfin l'obligation de le transmettre aux générations futures¹⁶⁷. Dans son introduction au recueil *Émotions patrimoniales*, Daniel Fabre¹⁶⁸ propose pour sa part cinq étapes, reprises d'un texte de Christian Jacquelin en 1993 : désigner, classifier, conserver, restaurer et publiciser. Dans *Heritage and digital humanities*¹⁶⁹, B. Dufrene s'appuie sur une référence à Jean-Michel Leniaud pour tenter d'étendre au patrimoine numérique les quatre éléments constitutifs du paradigme du patrimoine que sont les critères de conservation, les motivations pour accepter ou refuser le patrimoine, les manières dont il est géré et transmis, et les activités de dissémination. Christophe Gauthier, explorant le processus de patrimonialisation du cinéma¹⁷⁰ à l'aune des travaux de Fabre et Davallon, retient cinq phases : l'essor économique d'un nouvel objet, l'artification, l'institutionnalisation, l'intervention publique portant notamment sur la conservation, et enfin, l'émergence d'un marché de l'ancien.

Nous ne prétendons pas ici proposer un tour d'horizon exhaustif de tous les modèles, ni épuiser le détail de leurs différences ; cela serait d'autant moins pertinent que ces modèles théoriques ont été élaborés pour des objets différents du nôtre et ne peuvent donc lui être appliqués de manière directe. Ils peuvent cependant apporter l'indice de motifs récurrents que nous pourrions tenter d'utiliser comme clefs de lecture. Ainsi, le processus de patrimonialisa-

167 *Op. cit.*, p. 119 et ss.

168 Daniel Fabre, « Le patrimoine porté par l'émotion. » *op. cit.*

169 *Op. cit.* p. 11-28.

170 Article à paraître dans *Patrimoine et patrimonialisation du cinéma*, en ligne : <https://hal-enc.archives-ouvertes.fr/hal-02193020/>.

tion semble toujours naître d'une rupture, d'une découverte ou à tout le moins, de la prise de conscience d'un sentiment de perte. Il fait ensuite entrer en jeu un certain nombre d'acteurs dans une opération visant à désigner, authentifier, décrire, classer, enfin établir d'une manière ou d'une autre la nature patrimoniale de l'objet. La problématique de la conservation et des différents obstacles auxquels elle se heurte est ensuite posée. Enfin, la valorisation de l'objet, sa rencontre avec un public, est une étape indispensable du processus. Ce sont ces quatre angles, axes ou étapes, que nous avons retenus pour étudier les collections numériques au prisme d'un processus de patrimonialisation.

A. Une rupture ou une perte ?

La « trouvaille » de l'objet qui va devenir patrimonial initie, selon Jean Davallon, le processus de patrimonialisation. Il faut entendre ce terme de trouvaille au sens large d'une découverte, qui peut aussi découler d'un simple changement de regard : une ancienne usine que l'on souhaite protéger au titre du patrimoine industriel, par exemple, devient une « trouvaille » au moment de la prise de conscience de la perte que représenterait – pour le paysage, pour l'histoire de la ville, pour la mémoire de ses habitants – sa disparition. La notion de « trouvaille » est donc indissociable de l'idée, fondatrice de la notion de patrimoine, de rupture ou de prise de conscience d'une perte¹⁷¹.

Nous avons vu que dans le cas des bibliothèques et plus spécifiquement, du dépôt légal, la relation à la trouvaille est cependant très différente : ici, la patrimonialisation institutionnelle cherche à anticiper, par son exhaustivité, la possibilité de la trouvaille. On pourrait étendre cette réflexion aux archives, qui du point de vue de leur définition légale acquièrent ce statut dès l'instant où elles sont produites ; c'est ensuite l'archiviste qui, procédant à des opérations de gestion des fonds qu'il a collectés, effectue une sélection au sein de ces documents pour identifier ceux qui sont susceptibles d'avoir un intérêt patrimonial, et préserver l'opportunité de la trouvaille pour les générations futures. Ainsi, en archives comme en bibliothèque, le professionnel n'est pas l'auteur de la trouvaille : il prend en charge la préservation d'un champ plus large, par anticipation sur la perte ou disparition potentielle des objets concernés.

171 Michel Colardelle assimile même cette idée de perte au processus du deuil, dans Jacques Le Goff, *op. cit.* p. 127.

Dans le domaine du numérique, cependant, le sentiment de perte est réel, comme en témoignent les incessantes références aux problèmes de conservation que l'on rencontre dans la littérature professionnelle¹⁷². Jusqu'au début des années 2000, on espérait que des supports comme les « Century Disc » éviteraient la perte soudaine et irrémédiable des milliers de pages alors déjà numérisées¹⁷³. La volatilité du web, ressentie avec acuité par ceux qui tentent de l'utiliser au quotidien, a conduit à la même époque à l'émergence du souci de son archivage. L'analyse de gestion des risques, conduite en 2004 par un bibliothécaire australien, Gérard Clifton, en stage « profession culture » à la BnF, a joué un rôle moteur dans la mise en place du système SPAR¹⁷⁴.

La rhétorique de la perte fait ainsi partie intégrante de la construction des collections numériques, dès l'origine. Elle contribue également à légitimer les nouveaux objets patrimoniaux, en particulier dans le cas des archives de l'Internet : on compte chaque année le nombre de domaines qui ont disparu de la collecte large, et les parcours guidés fourmillent d'exemples de sites disparus, du web militant aux carnets d'écrivains.

B. Désigner, classer, décrire, fouiller ?

Dans le champ des monuments historiques ou de l'Inventaire, le « classement » a un sens bien différent de celui généralement employé en bibliothèque : dans le premier cas, il s'agit d'apposer sur l'objet une labellisation officielle qui l'érige officiellement au rang de patrimoine, alors que dans le second c'est l'organisation de la collection, et donc des documents les uns par rapport aux autres, qui se joue. Daniel Fabre emploie d'ailleurs le terme de « classier » pour désigner cette opération qui englobe inventaire, recensement et différents biais de reconnaissance comme la commémoration.

Il n'en reste pas moins que le patrimoine documentaire en général, et le dépôt légal en particulier, est étroitement associé à l'idée d'inventaire, de signalement, de description. La *Bibliographie nationale française* recense ainsi chaque année depuis 1811 l'ensemble des publications entrées par dépôt légal à la BnF¹⁷⁵. L'impossibilité de cataloguer les sites web faisant l'objet d'une collecte automatique, entérinée par le décret d'application correspondant qui admet que la bibliographie sera remplacée par un simple accès par URL, constitue, du

172 Voir ci-dessus, en première partie.

173 Jacqueline Sanson et Daniel Renoult (dir.), *La bibliothèque nationale de France... op. cit.* p. 146.

174 *Digital Preservation at the National Library of Australia*. Présentation de Gérard Clifton, stagiaire Profession Culture, 18 mai 2004. Archives numériques de la BnF : BnF-ADM-2008-024792.

175 Cf. <http://bibliographienationale.bnf.fr/Bicentenaire.html>.

moins en apparence, un premier accroc à ce principe. Le patrimoine numérique échapperait-il donc à cette obligation ? La situation n'est en réalité pas aussi simple. Nous l'avons vu, le signalement systématique du dépôt légal faisait déjà l'objet d'un certain nombre d'aménagements. *A contrario*, certaines ressources comme la presse quotidienne régionale numérique, tout en étant collectées par robot, font l'objet d'un bulletinage et d'un catalogage aussi complets que tout autre titre de presse. Certains départements de collection (notamment le département littérature et arts) ont fait le choix de créer dans le catalogue général les notices bibliographiques de centaines de sites web archivés dans le champ de leur discipline. La filière Acquisitions et Dons de Documents Numériques (ADDN), devenue opérationnelle en 2016, repose sur le principe d'un catalogage à la pièce au moment de l'acquisition. Il en est de même des livres numériques acquis à l'unité de manière pérenne. L'idée que le numérique n'a pas la possibilité ou la nécessité de bénéficier d'une description aussi précise que d'autres types de documents est donc démentie par la pratique.

Pour aller plus loin, il nous faut revenir sur la relation organique entre catalogue et collection qui existe en bibliothèque. Un document qui n'est pas catalogué, en bibliothèque, est un document qui n'est pas accessible : ceci vaut tant pour les collections traditionnelles que pour les collections numériques ; on peut aller jusqu'à dire que c'est le catalogue qui fonde l'existence de la collection¹⁷⁶. Or, dans le cas des collections numériques, le potentiel d'exploration et d'analyse ouvert par les technologies de fouille de texte et de données, ou tout simplement, la possibilité d'extraire des métadonnées à partir des documents eux-mêmes, fait émerger une nouvelle problématique : celle de l'apparition d'objets patrimoniaux nouveaux. Les articles au sein d'une revue, les plages dans un album de musique, un sujet de deux minutes au sein d'un journal télévisé deviennent des objets à part entière à partir du moment où ils bénéficient de leur propre description et peuvent donc être consultés en propre. L'application de procédures de fouille aux collections numériques, si elle débouche sur l'enrichissement des données d'accès aux collections, porte ainsi en gestation la démultiplication potentielle des trouvailles patrimoniales, à un niveau de granularité plus fin que ne le permettait le traitement des collections traditionnelles. C'est la nature autodescriptive des collections numériques, associée à l'existence et la maîtrise d'outils de traitements, qui conduit alors à un enrichissement considérable du patrimoine, non dans sa volumétrie mais dans la finesse de la connaissance qu'il est possible d'en avoir.

176 Cf. Gautier Poupeau, « Histoire(s) de notices. », *op. cit.*

C. Archive, mémoire, trace : quelle obligation de conserver ?

La conservation du patrimoine présente par essence, pour Daniel Fabre, un paradoxe : la vision occidentale s'attache à la conservation du patrimoine en sa matière, et non seulement en sa forme. Or, le numérique illustre particulièrement bien l'anecdote du bateau de Thésée, dont chaque pièce fut successivement remplacée jusqu'à ce qu'il n'en existe plus une seule qui soit d'origine. La préservation des documents numériques exige un constant renouvellement des supports, voire une migration des formats ou encore l'émulation des environnements logiciels d'origine. Les stratégies de pérennisation déployées par la BnF dans le cadre de SPAR reposent ainsi sur le concept de « cycle de vie » des paquets d'information¹⁷⁷ et celui de pérennisation de l'accès, qui s'oppose à l'idée de conservation pure et simple. Pourtant, cette approche ne fait pas l'unanimité : ainsi l'association MO5.com, consacrée à la préservation du patrimoine vidéoludique, défend-elle l'idée que les jeux vidéo doivent non seulement rester « jouables » mais encore, doivent l'être sur le matériel d'origine¹⁷⁸ : une position qui, malgré les précautions prises pour disposer de plusieurs exemplaires des jeux et appareils de lecture, conduit *in fine* à la détérioration des objets sous l'effet de l'usage.

À l'inverse, on pourrait également argumenter que le fait de « rejouer » les documents numériques patrimoniaux dans toutes leurs composantes – idée qui vaut pour le jeu vidéo mais aussi pour le web – n'est pas absolument nécessaire pour les appréhender d'un point de vue scientifique. Les historiens du web mobilisent ainsi de nombreuses sources secondaires pour approcher leur objet d'étude : ouvrages, modes d'emplois, magazines contemporains, archives orales, vidéos, copies d'écran... Les environnements en ligne hautement interactifs, comme le jeu *Second life* qui a eu un grand succès dans les années 2000, n'ont pas pu être archivés en tant que tels, ce qui n'empêche pas de les étudier à travers des sources secondaires : tel était par exemple le propos du projet « Preserving virtual worlds »¹⁷⁹. Une telle approche revient à considérer que les objets numériques en tant qu'artefacts techniques ne relèveraient pas en eux-mêmes de la conservation au sens documentaire, mais que leur patrimonialisation dépendrait de notre capacité à en conserver des traces permettant de les appréhender dans leurs différentes dimensions. Ainsi le dépôt légal, parce qu'il s'étend à des res-

177 Bertrand Caron, Jordan de la Houssaye, Thomas Ledoux et Stéphane Reecht, « Life and Death of an Information Package... » *op. cit.*

178 B. Barbier, *op. cit.*, p. 206.

179 Henry Lowood, « Memento mundi : are virtual worlds history ? » Présenté à IPRES 2009, the 6th International Conference on the Preservation of digital objects Proceedings, Oct. 2009, San Francisco. En ligne : <https://escholarship.org/uc/item/2gs3p6jx>.

sources documentaires triviales à l'origine – comme par exemple les magazines hebdomadaires ou mensuels sur l'informatique, ou le site Minitel.fr collecté en 1996 – recèle-t-il en lui-même des « trouvailles » susceptibles de permettre aux générations futures de comprendre ce qu'était la vie numérique des générations précédentes.

Le patrimoine numérique interroge donc fortement la notion d'« obligation de conserver » telle que la définit Jean Davallon. De multiples questionnements émergent : l'authentification elle-même, liée à un autre « geste » de patrimonialisation identifié par Davallon, est complexifiée par le fait que la pérennisation de l'accès aux objets numériques peut impliquer au moins leur copie sur un nouveau support et éventuellement leur transformation complète. Certaines de ces problématiques pourraient être rapprochées des enjeux que Daniel Fabre soulève s'agissant d'une autre étape, celle de la restauration. Il souligne en effet que la restauration « *a l'énorme défaut de choisir une version et d'imposer autoritairement une unique perception de l'œuvre ancienne alors que celle-ci doit rester ce qu'elle est devenue* » : en s'opposant à l'idée d'émulation des jeux vidéo, l'association MO5.com réfute elle aussi, d'une certaine manière, l'idée d'imposer une « perception unique » de l'œuvre, tronquée par l'émulation qui restitue tous les jeux dans un environnement et avec un matériel unique, alors que les différentes consoles d'origine présentaient des expériences sensorielles différentes. En tout état de cause, la transmission dans un bon état de conservation de l'objet issu du passé doit dans le cas du numérique céder devant une autre priorité : la pérennisation de l'accès. Cette situation est en partie compensée par la conservation de traces, non seulement de l'objet lui-même mais de son contexte de lecture¹⁸⁰.

D. Transmettre, montrer : peut-on exposer le patrimoine numérique ?

Exposition, publication, mise en accès public d'un monument, valorisation sont différentes formes de mise en relation de l'objet patrimonial avec son public, indispensables – quel que soit le modèle – pour compléter le processus de patrimonialisation. Pour Jean Davallon, l'exposition joue un rôle essentiel à plusieurs niveaux : elle célèbre la trouvaille, entérine le rôle de représentation de l'objet vis-à-vis de son monde d'origine, et concrétise sa fonction de lien entre le passé, qu'il représente, le présent qui le célèbre et le futur auquel il sera transmis. Dans cette logique, l'idée de patrimoine numérique ne peut se concevoir que si elle embrasse

180 À la BnF, la collection Charles Cros a ainsi pour objectif de conserver des exemples de machines de lecture de toutes sortes. Cf. Marianne Deraze et Xavier Loyant, « Quand la collection Charles Cros s'expose en ligne. » Dans *Bulletin de l'AFAS* n° 42, 2016. En ligne : <http://journals.openedition.org/afas/2983>.

une composante de mise en relation de ce patrimoine avec un public qui ne se limite pas aux chercheurs fréquentant les salles de lecture du Rez-de-jardin de la BnF. Le patrimoine écrit, si on le considère dans sa composante traditionnelle de document « ancien, rare et précieux », n'échappe pas non plus à cette logique, car les documents entrant dans cette catégorie sont justement ceux que l'on est susceptible d'exposer, que le propos soit historique, scientifique, esthétique ou commémoratif¹⁸¹.

Les expositions consacrées à des objets numériques, sans être monnaie courante, ne sont pas inexistantes : deux d'entre elles ont présenté le patrimoine vidéoludique à Paris, *Museo-games* en 2010 au Conservatoire national des Arts et Métiers, et *Game Story* en 2011 au Grand Palais¹⁸². D'autres événements consacrés à l'art numérique ont eu lieu au Centre Pompidou : on mentionnera notamment, en 2012, l'exposition *Multiversités créatives* qui présentait des travaux sur les données, des impressions 3D, etc. La participation de la BnF à la *Paris Games Week* ou encore l'organisation de colloques et journées d'études sur le thème du patrimoine vidéoludique ont contribué à asseoir la légitimité de ces objets en tant que collections de la bibliothèque et patrimoine national. Plus largement, le document multimédia, qu'il soit audiovisuel numérisé ou nativement numérique, trouve volontiers sa place dans les expositions. Le numérique est par ailleurs souvent employé comme outil de médiation : application accompagnant l'exposition, projection, feuillettoir sur tablette permettant de consulter en intégralité le manuscrit exposé sous vitrine, exposition virtuelle faisant écho à l'événement...

Cependant, ces différentes manifestations du numérique dans l'espace de monstration physique de la culture où celle-ci rencontre son public semblent relativement anecdotiques ou en tout cas, insuffisantes à illustrer pour le numérique le processus expliqué par Jean Davallon dans toute sa complexité symbolique. Dans le présent, l'objet en tant que représentant de son monde a une fonction sociale. Celui qui visite l'exposition en devient un bénéficiaire, non pas propriétaire, mais plutôt dépositaire de ce rôle symbolique avec la responsabilité de transmettre à son tour l'objet aux générations futures¹⁸³. Pour observer un tel phénomène dans un contexte numérique, sans doute ne faut-il pas se limiter à l'espace de monstration physique, mais également analyser le lien social qui se tisse autour du patrimoine numérique dans le cadre du réseau.

181 À ce titre, on peut citer l'exposition *Manuscrits de l'extrême* en 2019 à la BnF, qui a présenté nombre d'objets inattendus, parfois éloignés de ce que l'on imagine habituellement quand on parle de manuscrits.

182 B. Barbier, *op. cit.*, p. 8.

183 Jean Davallon, *op. cit.*, p. 123-125.

À l'occasion de la commémoration du centenaire de la Grande Guerre, entre 2014 et 2018, le projet de recherche « Le devenir en ligne du patrimoine numérisé de la Grande Guerre » s'est justement penché sur la relation des réseaux d'amateurs au patrimoine numérisé et mis en ligne par des institutions comme la BnF ou la BDIC¹⁸⁴, montrant que le web crée une valeur d'échange des documents numérisés qui vient s'ajouter à leur valeur d'usage¹⁸⁵. La cartographie des sites privés, associatifs et institutionnels consacrés à la Grande Guerre fait apparaître des communautés d'amateurs fortement connectées, qui s'approprient les collections numérisées mises en ligne par les bibliothèques. L'enquête qualitative par entretiens conduite au cours du projet montre en outre que la démocratisation qui découle de la mise en ligne de bibliothèques numériques comme Gallica, accessibles à tous gratuitement et sans formalité, constitue un véritable tournant : beaucoup de ces « amateurs », passionnés de la Grande Guerre et contributeurs actifs de l'élaboration en ligne de nouveaux « lieux de mémoire numériques »¹⁸⁶, n'auraient pas pu ou pas souhaité entreprendre de telles activités sans cet accès inespéré au patrimoine numérisé. Percevant Gallica comme un vaste réservoir sans véritable inventaire, ils se consacrent à un dépouillement laborieux dans l'espoir de la fameuse « trouvaille » : médiateurs des ressources qu'ils ont alors pu identifier, ils deviennent acteurs de leur patrimonialisation, de la transformation de la mémoire en histoire, et donc de cette transmission aux générations futures.

Au-delà de ces « amateurs », fortement investis dans cette tâche de transmission, la conquête des réseaux sociaux entreprise par la BnF annoncée dans l'initiative *Numérique horizon 2015*¹⁸⁷ apparaît comme un moyen de toucher le « grand public » et donc d'élargir encore davantage le champ des bénéficiaires du patrimoine numérique. Les formes de valorisation ludiques comme le jeu du vendredi sur Facebook ou la « chasse au trésor » du lundi matin¹⁸⁸ semblent avoir joué, dans un monde numérique, un rôle qui rappelle également le principe de la « trouvaille » de Davallon, étendu à toute une communauté en ligne. Fortement liée à l'événementiel et au quotidien, avec des opérations comme le Tour de France de Gallica durant l'été 2017 ou la valorisation du manuscrit de *Notre-Dame de Paris* de Victor Hugo en

184 Bibliothèque de Documentation et d'Information Contemporaine, aujourd'hui appelée la Contemporaine.

185 Muriel Amar et Philippe Chevallier dans *Le web français de la Grande Guerre... op. cit.* p. 22.

186 Valérie Baudouin dans *Le web français de la Grande Guerre... op. cit.* p. 58-59.

187 Voir deuxième partie ci-dessus.

188 Les nouveaux documents numérisés étant mis en ligne chaque semaine le lundi, il est devenu une tradition au sein de la communauté des Gallicanautes d'y chercher des pièces remarquables pour les partager sur Twitter.

mai 2019 juste après l'incendie de la cathédrale¹⁸⁹, ce type de médiation est susceptible de toucher le plus grand nombre et remplit bel et bien la fonction de lien entre le passé et le présent qui était celle de l'exposition d'après Davallon. Les réseaux sociaux, avec leurs fonctions de commentaires, de partages, de « j'aime », conduisent naturellement à l'amplification du phénomène, jusqu'à la viralité dans le cas des opérations les plus réussies.

189 Respectivement, dans *Le blog Gallica* : François Michel, « Le tour de France de Gallica : prêts pour le départ ! » 27 juin 2017, en ligne : <https://gallica.bnf.fr/blog/27062017/le-tour-de-france-de-gallica-pret-pour-le-depart> et Thomas Cazentre, « Le manuscrit de Notre-Dame de Paris. » 20 mai 2019, en ligne : <https://gallica.bnf.fr/blog/20052019/le-manuscrit-de-notre-dame-de-paris>.

III. Émotions et patrimoine à la lumière de la culture numérique

Le 5 janvier 2017, un message sur le compte Twitter de Gallica invite les internautes à un « Tour de France des spécialités culinaires », à travers une carte de 1929 qui détaille la gastronomie de chaque région¹⁹⁰. Rapidement relayée par plusieurs sites de presse régionale¹⁹¹, la *Carte gastronomique de la France d'Alain Bourguignon*¹⁹² connaîtra un succès extraordinaire, souligné par Adrien Nouvellet dans son étude des traces d'usage de Gallica¹⁹³. Chacun s'amuse à zoomer et republie à son propre compte la spécialité locale de son enfance, contribuant à la diffusion de ce document patrimonial qui était pourtant en ligne, au sein de la masse de la bibliothèque numérique, depuis plusieurs années¹⁹⁴. Hasard de calendrier, amplifié par ce relai médiatique inhabituel ? Engouement pour le patrimoine immatériel, ici gastronomique, des territoires ? On peut aussi lire dans cette anecdote la concrétisation dans un monde numérique des « émotions patrimoniales » que Daniel Fabre analyse dans l'ouvrage éponyme paru en 2012. Ces émotions sont au nombre de quatre, sur une échelle allant de la plus tiède à la plus ardente : le transport, la dispute, la déploration et la sédition¹⁹⁵. Intervenant à toutes les étapes du processus de patrimonialisation, elles caractérisent l'« âge du patrimoine » où la mise « en valeur », au sens littéral du terme, des objets patrimoniaux n'est plus le fait d'un acte régalien mais exprime l'irruption d'un sentiment démocratique du passé ; où la relation aussi bien individuelle que collective au patrimoine est bouleversée.

Les détracteurs du numérique, au tournant des années 2000, étaient convaincus et tenaient pour acquis qu'aucune émotion ne pouvait être associée au support froid et sans âme d'un écran¹⁹⁶. Pensée qui conduit à remettre en cause l'idée de la transmission générationnelle du patrimoine : face à l'immédiateté, l'ubiquité, la volatilité du web, l'« obligation de conserver » identifiée par Jean Davallon comme le dernier des gestes du processus de patrimoniali-

190 <https://mobile.twitter.com/GallicaBnF/status/816930189842251776>.

191 *Sud Ouest* le 6 janvier : <https://www.sudouest.fr/2017/01/06/sud-ouest-une-carte-des-specialites-gastronomiques-et-des-vignobles-de-1929-en-version-numerique-3083116-708.php>, *France 3 Bretagne* le 12 janvier : <https://france3-regions.francetvinfo.fr/bretagne/carte-quelle-specialite-culinaire-votre-commune-1929-1173095.html>.

192 <https://gallica.bnf.fr/ark:/12148/btv1b52504043q>

193 Adrien Nouvellet, Valérie Beaudouin *et al.*, *op. cit.*

194 Cf. Thomas Perrono et Erwan Le Gall, « Le dessous des cartes... gastronomiques. A propos d'un étonnant buzz sur twitter lancé par Gallica. » Dans *En envor*, s.d. En ligne : http://enenvor.fr/eo_actu/entredouxguerres/le-dessous-des-cartes-gastronomiques-a-propos-d-un-etonnant-buzz-sur-twitter-lance-par-gallica.html.

195 Daniel Fabre, *op. cit.* pp. 37-38.

196 F. Stasse écrit ainsi en 2002 : « à l'évidence, aucun de ces sentiments ne supporterait d'être exprimé sur un écran d'ordinateur par l'entremise glaciale des livres numérisés » (*op. cit.*, p. 46).

sation semble devenir une mission impossible¹⁹⁷. Pourtant, les quinze dernières années ont vu s'exprimer, dans le champ du numérique, ces émotions qui caractérisent l'ère d'une appropriation par les communautés de la notion de patrimoine. Nostalgie, expression d'un attachement intime, protestation, déploration, contribution et participation, jusqu'à la définition de « communs » qui échappent aux lois du marché : autant d'expressions individuelles et collectives vis à vis de ce nouveau patrimoine numérique constitué par la BnF, qui témoignent d'un profond sentiment d'appartenance et de l'émergence d'une culture nouvelle, une culture numérique.

A. Une culture numérique ? (le transport)

Si le patrimoine émerge de sa reconnaissance par une communauté, parler de patrimoine numérique implique de s'interroger sur le « monde » qu'il représente, au sens de Jean Davallon : celui de la culture numérique. Pour Milad Doueïhi¹⁹⁸, c'est le partage, c'est-à-dire l'émergence d'une valeur économique et sociale, qui constitue l'indice de la naissance de cette culture numérique. Dans les sociétés occidentales, le numérique est devenu indissociable de presque toutes les activités humaines, que l'on s'intéresse au plan documentaire, juridique, politique, économique, social, etc. On peut dès lors parler d'impact culturel global, au sens large du mot culture qui englobe toutes les pratiques, connaissances, normes et traditions qui sous-tendent globalement le fonctionnement de la société¹⁹⁹.

Vu comme une culture qui modifie notre rapport au monde, le numérique embrasse plusieurs dimensions qui touchent au quotidien et à l'intime. L'irruption de la micro-informatique dans les foyers et les bureaux depuis les années 1980 trouve un aboutissement dans les années 2000, avec le succès de l'iPhone puis de l'iPad, puis avec les objets connectés qui la rendent omniprésente et la rapprochent de plus en plus du corps. S'y ajoute la connectivité permanente qui transforme en profondeur la relation de chacun à l'information qui circule sur les réseaux mais aussi aux autres personnes qui y sont connectées, modifiant profondément l'espace-temps des interactions sociales. En 2012, Michel Serres invente le personnage

197 Régis Debray s'inquiète ainsi, dans une conférence prononcée en décembre 2000 à la BnF : « Nous sommes sans doute la première culture à qui ses équipements techniques permettent de fuir ses devoirs envers les générations passées et futures dans ses impératifs immédiats de circulation. » Dans *Les diagonales du médiologue*, Paris : BnF, 2001. En ligne : <https://books.openedition.org/editionsbnf/1121?lang=fr>.

198 *Qu'est-ce que le numérique ?* Paris : Presses Universitaires de France, 2013. En ligne : <https://www.cairn.info/qu-est-ce-que-le-numerique--9782130627180-page-5.htm>.

199 Cf. Emmanuelle Bermès, « Qu'est-ce que le numérique ? » dans *Figoblog*, 2 août 2019. En ligne : <https://figoblog.org/2019/08/02/quest-ce-que-le-numerique/>.

de *Petite poucette*²⁰⁰ pour caractériser la nouvelle génération, familière de cet environnement numérique dans lequel la relation à la connaissance est inversée. Le choix de la référence au pouce, le doigt de la main qui est le plus impliqué dans la manipulation du smartphone, évoque le devenir de la notion de « digital » : le numérique a bien une dimension physique, il devient une extension du corps humain et entraîne une profonde mutation des facultés cognitives humaines, notamment la mémoire.

De la mémoire au patrimoine, il n'y a qu'un pas. La première émotion patrimoniale définie par Daniel Fabre, le « transport » – au sens littéraire et suranné d'élan de joie – émerge de l'adhésion des individus à l'appréciation d'un patrimoine qui leur est à la fois remarquable et familier. L'illustration la plus emblématique en est la promenade « le nez en l'air » dans les rues des villes et villages, où l'on découvre les richesses architecturales les plus ordinaires aussi bien que les plus imposantes, et à l'issue de laquelle on partage ses coups de cœur : un grand « j'aime » qui n'a plus besoin de fondement historique ou de reconnaissance scientifique pour être exprimé par tout un chacun. Quand Daniel Fabre écrit ces mots en 2012, songe-t-il également au « j'aime » popularisé par Facebook, qui règnera bientôt en maître sur les réseaux sociaux²⁰¹ ? Ce n'est pas certain, et pourtant, il est évident que les facilités de communication apportées par le web d'abord, puis par les réseaux sociaux ont considérablement servi le développement de ces communautés qui se donnent pour mission d'élire leur propre patrimoine²⁰².

Le patrimoine numérisé, notamment à travers sa circulation sur les réseaux sociaux, devient naturellement l'instrument de cette émotion de transport individuel et collectif qui s'exprime par l'admiration partagée. Les responsables de l'équipe réseaux sociaux de Gallica remarquent ainsi que les images de manuscrits médiévaux ont systématiquement un grand succès quand elles sont partagées en ligne²⁰³ : un large public est sensible au privilège qui lui est accordé de voir ces trésors, habituellement réservés aux chercheurs dans le silence d'une salle de lecture à l'accès très restreint ou aux vitrines d'une exposition. Dans d'autres cas, comme celui de la *Carte gastronomique de la France* cité plus haut ou encore de l'opération

200 Michel Serres, *Petite poucette*. Paris : Le Pommier, 2012.

201 Sur Twitter, l'étoile qui servait à mettre une publication en favori a été remplacée par un cœur en 2015.

202 Arnaud Dhermy et Valérie Tesnière, « Les sociétés savantes, des revues au web. » Dans *Le web français de la Grande Guerre... op. cit.* p. 143.

203 Cf. Tuija Ainonen, Thomas Henry et Claire Séguret, « Opening Medieval Manuscripts to a Larger Audience: Challenge of Social Media Culture. » Présentation pour la conférence France et Angleterre : manuscrits médiévaux entre 700 et 1200. Paris, 21-23 novembre 2018.

« mon bled dans Cassini » en 2015²⁰⁴, le patrimoine numérisé est prétexte à une plongée dans l'histoire locale et personnelle de chacun. C'est ainsi une relation individuelle, intime au patrimoine qui se crée par l'intermédiaire du numérique. Mais pour aller plus loin, il faut également considérer ce qu'il en est du patrimoine nativement numérique.

Si Petite Poucette voit sa relation au monde profondément transformée par l'omniprésence du numérique dans son environnement aussi bien corporel qu'émotionnel, elle se construit également dans ce contexte qui en vient à faire partie de sa mémoire personnelle. Des collectionneurs de jeux vidéo, fans de rétrogaming²⁰⁵, au portail des gifs animés réalisé à l'occasion des vingt ans des archives du web par Internet Archive²⁰⁶, une nostalgie du numérique commence à émerger à l'heure où celui-ci atteint l'âge d'une génération. L'idée que la BnF, au sein des milliards d'URL préservées dans SPAR au titre du dépôt légal, détient peut-être une copie du « skyblog honteux de vos années collègues »²⁰⁷ est ainsi emblématique de la relation des citoyens de l'État-Nation à leur bibliothèque nationale à l'heure du numérique. Comme l'a montré l'étude sur l'image de la BnF en 2017²⁰⁸, les Français ne se voient pas forcément comme des usagers ou des visiteurs de la BnF, mais ils se reconnaissent, avec émotion et étonnement, dans la mémoire de leur époque qu'élabore l'institution, y compris s'agissant du numérique.

B. Patrimoine numérique, sphère publique et acteurs privés (la dispute)

La dispute, émotion qui s'exprime par la polémique et le débat, peut survenir à chaque étape du processus de patrimonialisation. Elle lui est même intrinsèquement liée, car ériger un objet en patrimoine est un choix, et tout choix doit être controversé²⁰⁹. Faut-il dérestaurer les ajouts de Viollet-le-Duc à la basilique St-Sernin de Toulouse ? Est-il légitime de fermer la grotte de Lascaux pour la préserver, et dès lors ne proposer aux yeux du public qu'un fac-similé ? La flèche de Notre-Dame doit-elle être reconstruite à l'identique ? Ce qui caractérise « l'âge du patrimoine », c'est d'une part la propension de ces polémiques à gagner l'espace

204 Claire Chemel et Louis Jaubertie, « La Carte générale de la France de Cassini. » Dans *Le blog Gallica*, 5 juin 2015. En ligne : <https://gallica.bnf.fr/blog/05062015/la-carte-generale-de-la-france-de-cassini-0>.

205 B. Barbier, *op. cit.* p. 121.

206 <https://gifcities.org/>

207 Annabelle Laurent, « Le Skyblog de vos années collègues est (peut-être) archivé à la BnF. » Dans *20 minutes*, 25 novembre 2016. En ligne : <https://www.20minutes.fr/culture/1967703-20161125-skyblog-honteux-annees-college-peut-etre-archiver-bnf>.

208 Thierry Pardé, *Résumé exécutif de l'étude sur l'image de la BnF auprès de la population française*. 20 août 2018. Archives numériques de la BnF : BnF-ADM-2018-094200.

209 Daniel Fabre, *op. cit.* p. 70.

public, bien au-delà des cercles des spécialistes, et d'autre part la remise en cause des choix effectués par l'autorité en charge de ce patrimoine, que celle-ci soit politique, scientifique ou professionnelle.

Le numérique, en particulier lorsqu'il s'agit de la BnF, ne fait pas exception. Nous avons rappelé que dès l'origine du projet, en 1988, la nouvelle bibliothèque portait en son sein la perspective d'un chantier de numérisation considérable, assorti d'une ambition de large diffusion qui a fini par voir le jour, même si l'irruption du web l'a profondément transformé. Objet de polémiques en tout genre, la très grande bibliothèque voit avant 2000 les médias s'empoigner sur le sujet du projet d'architecture, la question de la césure des collections²¹⁰ ou encore les problèmes du système informatique. Cependant, c'est avec Jean-Noël Jeanneney en 2005, et « grâce », d'une certaine manière, à l'arrivée sur la place du débat d'un acteur privé, Google, que la numérisation des collections de la BnF devient un sujet de dispute.

Si l'opposition jusqu'au-boutiste de Jean-Noël Jeanneney contre Google suscita de nombreuses critiques, on peut en dire tout autant de la position plus mesurée de Bruno Racine qui cherche, en 2013, une voie de complémentarité entre sphère publique et financeurs privés. Une filiale de la BnF, structure de droit privée nommée BnF-Partenariats, est alors créée pour permettre à l'institution d'émarger au dispositif des investissements d'avenir, ou « grand emprunt ». Sa mission est de combiner ces financements avec des apports issus du privé pour conduire de nouveaux projets de numérisation avec un modèle économique et un retour sur investissement. Le partenariat signé avec l'éditeur Proquest, qui porte sur la numérisation des ouvrages de la réserve des livres rares, fait grand bruit : on reproche à la BnF la durée, jugée excessive, pendant laquelle cette numérisation fait l'objet d'une exclusivité de diffusion pour le partenaire privé, interdisant la diffusion dans Gallica de ces numérisations²¹¹. Ce sont alors principalement les associations professionnelles qui questionnent d'un point de vue déontologique la démarche de la BnF et sa conformité avec la recommandation du Comité des Sages européen de 2011²¹².

L'autre projet qui mérite d'être mentionné au titre de la dispute dans la sphère du patrimoine numérique touche également à la question de la relation entre acteurs publics et privés :

210 Claudie Voisenat, « Les jeux de l'émotion et de la raison : la polémique autour des collections de la Bibliothèque nationale (1988-1992). » Dans Daniel Fabre (dir.), *op. cit.* p. 213-236.

211 Eric Chaverou, « Numérisation polémique à la BnF. » Site de France Culture, 5 février 2013. En ligne : <https://www.franceculture.fr/economie/numerisation-polemique-la-bnf>.

212 *Stratégie numérique: le Comité des Sages escompte une «nouvelle Renaissance» grâce à la mise en ligne du patrimoine culturel européen.* Communiqué de presse de la Commission Européenne, 10 janvier 2011. En ligne : https://europa.eu/rapid/press-release_IP-11-17_fr.htm.

il s'agit du projet ReLire, qui vise à numériser puis remettre en vente sous forme numérique des ouvrages indisponibles du vingtième siècle. Partie-prenante au titre de la numérisation, qui doit s'effectuer sur les collections patrimoniales du dépôt légal, la BnF se voit également confier le rôle d'opérateur en charge du registre des œuvres indisponibles, à partir duquel les auteurs et ayants droit peuvent demander à sortir du dispositif de licence globale. C'est ce principe de l'« *opt-out* », fondamentalement incompatible avec la conception française du droit d'auteur, qui suscite l'ire cette fois d'un public bien plus large, emporté notamment par des auteurs comme François Bon²¹³. Le dispositif, qui vise la remise en vente des titres concernés et privilégie les éditeurs qui ont signé les contrats d'édition pour une publication papier, est vu comme favorisant une approche mercantile de la culture. L'État comme la BnF en tant qu'opérateur sont pointés du doigt, malgré les négociations préalables et le processus législatif : ce qui est considéré par les uns comme une mission de service public – à savoir rendre à nouveau disponible, dans un dispositif respectueux des droits de propriété intellectuelle, un patrimoine enfoui dans les tréfonds des collections nationales – est vu par les autres comme une trahison envers les auteurs comme les lecteurs.

Il faudrait étudier plus en profondeur ces disputes qui agitent l'histoire de la numérisation à la BnF pendant dix ans pour mieux comprendre comment l'intervention conjointe d'acteurs publics et privés conduit à la revendication par différentes communautés – les professionnels des bibliothèques, les auteurs, les lecteurs et chercheurs – d'un nouveau patrimoine numérique dont les caractéristiques sont, à l'issue de ces débats, bien définies. Issu du domaine public, sa diffusion en ligne de façon ouverte et gratuite est devenue une obligation et une évidence. Par extension, l'idée que les archives du web, constituées majoritairement de documents qui étaient en accès ouvert au moment de leur collecte, ne soient consultables que par les chercheurs dans les emprises de la bibliothèque suscite également une forte incompréhension²¹⁴, même si droit d'auteur et droit à l'oubli²¹⁵ justifient clairement cette restriction. Le terme d'enclosure est employé pour désigner la création par la bibliothèque de nouvelles restrictions d'accès ou d'usage par rapport à des documents perçus comme faisant partie de

213 « Auteurs, contre l'État voleur, réclamez vos droits ! » Dans *Tiers Livre*, 1^{ère} mise en ligne 23 mars 2013 et dernière modification le 9 septembre 2013. En ligne : <https://www.tierslivre.net/spip/spip.php?article3447>.

214 Pointée, par exemple, par Michel Melot dans « Digital archives or anti-heritage ». Dans B. Dufrière, *op. cit.* p. 39-48.

215 Cf. **Emmanuelle Bermès et Clément Oury, « Web 2.0 et mémoire : de la conversation à la conservation ». Dans *Documentaliste - Sciences de l'information* vol. 46 n°1, 2009, pp. 61-63.

l'espace public²¹⁶. C'est alors le caractère de patrimoine qui prime et conduit à la revendication d'un droit d'accès de chaque citoyen à ces nouvelles collections numériques.

C. Le patrimoine numérique en actions (la déploration)

Fort heureusement, à ce jour la BnF n'a pas (encore ?) été confrontée à une catastrophe patrimoniale d'ampleur conduisant à la perte d'une partie conséquente des collections numériques. On peut en partie en attribuer le mérite aux mesures de conservations prises, notamment avec SPAR, car si des épisodes de problèmes techniques ont bien eu lieu, le fait de disposer de multiples copies a permis d'en limiter la portée, au moins sur le plan patrimonial, c'est-à-dire d'éviter les pertes de collections (l'impact économique n'étant cependant pas nul). Cependant, nous l'avons vu plus haut²¹⁷, le numérique – de par sa volatilité intrinsèque – est bel et bien fréquemment associé à un sentiment de perte, ce qui permet de vérifier l'enchaînement identifié par Fabre en cas de catastrophe patrimoniale : lamentation, action et reconstruction à l'identique.

La déploration est une émotion fondamentalement collective, qui conduit à un besoin d'unité de la communauté face au désastre, cette unité se réalisant dans l'action. La question de la mobilisation des communautés pour la préservation de leur propre patrimoine numérique doit donc être examinée. Dans le cas du jeu vidéo, qui se caractérise par l'existence d'une importante communauté d'amateurs ou de fans²¹⁸ fortement structurée en associations, ce phénomène est particulièrement facile à observer.

Il existe cependant une autre dimension dans laquelle intervient le patrimoine numérique, qui nous semble intéressante à considérer et qui est liée au concept de déploration, vu comme un catalyseur de l'unité d'une communauté face à la perte, potentielle ou effective, de son patrimoine. La numérisation a en effet, ces dernières années, été appelée à jouer un rôle prépondérant dans les débats autour de la restitution des œuvres culturelles²¹⁹. L'enjeu est de proposer via la numérisation un substitut à la restitution à leur communauté d'origine des œuvres aujourd'hui conservées dans des institutions occidentales, notamment suite à l'histoire

216 Lionel Maurel, « Commun vs enclosures : quand les bibliothèques publiques verrouillent l'accès aux biens communs informationnels. » Dans *BIBLIOTHÈQUE(S)*, 2014, n° 76, pp.15-21. En ligne : <https://hal-univ-paris10.archives-ouvertes.fr/hal-01373179/>.

217 Voir 3^e partie, II, a.

218 B. Barbier, *op. cit.* pp. 89 et ss.

219 De nombreuses polémiques ont émergé sur ce thème, notamment autour du rapport de Felwine Sarr et Bénédicte Savoy : *Rapport sur la restitution du patrimoine culturel africain. Vers une nouvelle éthique relationnelle*. Novembre 2018. En ligne : http://restitutionreport2018.com/sarr_savoy_fr.pdf.

coloniale des nations concernées. La section des bibliothèques nationales de l'IFLA a ainsi lancé en 2017 un groupe de travail sur l'« unification numérique », faisant suite à la *Recommandation concernant la préservation et l'accessibilité du patrimoine documentaire, y compris numérique* adoptée par l'UNESCO en 2015, en particulier sa dernière partie qui porte sur la coopération nationale et internationale. En 2019, ces travaux ont conduit à la publication d'un guide pour la mise en place d'un projet de réunification numérique²²⁰. À la BnF, la collection *Patrimoines partagés*, lancée en 2016, illustre ce type d'initiative avec des projets comme *Bibliothèques d'Orient*²²¹.

Le recours à la numérisation pour des projets visant à réunir le patrimoine écrit d'une même communauté, que les aléas de l'histoire ont dispersé, a pour conséquence l'émergence de projets dont l'impact international est bien supérieur à une simple restitution : il implique une coopération durable entre institutions, mobilise diverses sources de financement dont le mécénat, offre des opportunités de valorisation considérables que ce soit en ligne ou via l'organisation de colloques et autres événements... On pourrait ainsi parler de processus de repatrimonialisation par le numérique : les différentes étapes, allant de l'identification du patrimoine écrit d'une communauté à sa numérisation, puis sa mise à disposition d'un public dans un nouvel écran via le numérique, sont mises en œuvre. Elles mobilisent un cercle d'acteurs identifiés pour leur autorité scientifique ou institutionnelle. Le produit qui en est issu, ainsi « mis en valeur » au sens premier du terme, représente davantage que la somme des objets qu'il réunit : c'est un nouveau patrimoine réunifié qui émerge grâce au numérique, incarnant une communauté qui n'y avait plus accès à cause de sa dispersion. C'est la dimension « virtuelle » de la culture numérique qui s'exprime ici : au sens de la création de nouveaux mondes, qui sont le prolongement du monde sensible, mais aussi de l'extension des possibles grâce à une technologie fondée sur l'ubiquité et la réduction des barrières spatiales et temporelles.

D. Le patrimoine vu comme un bien commun (la sédition)

Dans un monde numérique, l'expression des émotions patrimoniales, du « j'aime » de Facebook à l'engagement dans des associations de sauvegarde du jeu vidéo comme MO5.com, prend une forme immédiatement contributive. Tout un chacun peut très facilement

220 *Guidelines for Setting Up a Digital Unification Project*. IFLA, 31 août 2019. En ligne : <https://www.ifla.org/publications/node/92435>.

221 <https://heritage.bnf.fr/bibliothequesorient/fr/homepage>.

devenir acteur du partage, de la dispute ou des actions entreprises pour conserver ou reconstituer un patrimoine menacé. Cette ère de la contribution, qui n'est pas spécifique au fait patrimonial, mais irrigue tous les aspects de la société numérique, est désormais bien caractérisée²²². Dans les bibliothèques, ce phénomène a suscité très tôt l'intérêt des professionnels, qui le voyaient comme un moyen de mobiliser les savoirs et le temps des usagers pour contribuer à la mission jamais achevée de description des collections – parfois à un niveau très fin. La BnF s'est ainsi engagée, entre 2012 et 2014, dans un projet avec France Télécom pour réaliser un outil collaboratif de correction de l'OCR. Le prototype Correct, mis en service en 2014, a fait l'objet d'une étude d'usage, démontrant que si les contributeurs étaient peu nombreux, les plus engagés et organisés revendiquaient un désir très explicite de contribuer à améliorer Gallica, bibliothèque numérique nationale envers laquelle ils exprimaient un fort sentiment d'appartenance²²³.

On peut citer de nombreuses initiatives dans le domaine du patrimoine des bibliothèques, qui ont par la suite proposé des approches participatives ou contributives à leurs usagers en ligne : transcription des menus archivés par la New York Public Library²²⁴, géolocalisation de photographies à la bibliothèque royale du Danemark²²⁵... En 2017, la British Library ouvre le site LibCrowds, consacré aux projets de *crowdsourcing*²²⁶. À la BnF, le site Gallica Studio²²⁷ est également mis en ligne fin 2017 dans l'objectif de valoriser les créations et contributions réalisées par les Gallicanautes autour de Gallica, en incluant les projets lauréats du hackathon lancé en 2016. Il ressort de ces différentes initiatives, certaines couronnées de succès, d'autres ignorées et délaissées, une interrogation fondamentale sur le positionnement de la bibliothèque face à l'appropriation du patrimoine numérique par les communautés²²⁸. La motivation essentielle de la contribution n'est pas, ne peut pas être d'aider l'institution à accomplir les tâches de traitement qui sont les siennes dans le domaine de la conservation et de l'accès au patrimoine. Elle relève plutôt d'un besoin d'appropriation de ce patrimoine par l'action : l'affirmation « *le patrimoine, c'est nous* » posée par Daniel Fabre n'a jamais été

222 Voir par exemple Patrice Flichy, *Le sacre de l'amateur : sociologie des passions ordinaires à l'ère numérique*. Paris : La République des idées / Seuil, 2010 ; et également Nicolas Colin et Henri Verdier, *L'âge de la multitude : entreprendre et gouverner après la révolution numérique*. Paris : Armand Colin, 2015.

223 Isabelle Josse, « Résultats de l'étude d'usages de la plateforme Correct. » *Blog Gallica*, 1^{er} juin 2015. En ligne : <https://c.bnf.fr/ExA>.

224 <http://menus.nypl.org/>

225 <http://www.kb.dk/danmarksetftraluftten/>

226 <https://www.libcrowds.com/>

227 <http://gallicastudio.bnf.fr/>

228 Céline Leclair, « La BnF à l'âge de la multitude ». Dans *Bulletin des bibliothèques de France*, 2017, n° 13, p. 142-143. En ligne : http://bbf.enssib.fr/matieres-a-penser/la-bnf-a-l-age-de-la-multitude_68032.

plus vraie que lorsque les citoyens sont mis en capacité, par l'ouverture des données, la création d'outils *ad hoc* et l'invitation de l'institution à s'impliquer dans des « chantiers », de créer eux-mêmes le patrimoine numérique – ou de le mettre en lumière en le faisant émerger d'une masse indistincte, ce qui revient au même.

Mais pour cela, il faut que la bibliothèque accepte une certaine externalité du patrimoine : elle doit renoncer à maîtriser de bout en bout le processus, accepter la légitimité des communautés à élire et traiter leur propre patrimoine, et trouver un juste rôle dans la valorisation de ce type d'action. Ce rôle ne peut être ni une caution institutionnelle aveugle à des contenus qui n'auraient pas suivi le processus plus officiel de qualification de l'information, ni un contrôle trop excessif, au risque de facilement prendre la coloration de la censure.

Lorsque l'action de la communauté remet en cause les règles définies par l'autorité scientifique, professionnelle ou politique incarnée par l'institution, elle peut aller jusqu'à employer des moyens d'actions hors de la légalité : on touche alors à la sédition, dernière émotion, la plus extrême, identifiée par Daniel Fabre. Celui-ci reconnaît que la sédition est un fait relativement rare dans le contexte du patrimoine ; il en cite néanmoins quelques exemples²²⁹. Dans le domaine numérique, on pourrait invoquer le discours tenu par le responsable du site Abandonware²³⁰ qui reconnaît que la mise en ligne du code des anciens jeux vidéo « abandonnés », si elle est cadrée par un certain nombre de règles (notamment la vérification que le jeu n'est plus exploité commercialement et le retrait sur simple demande), reste totalement illégale.

Le militantisme contre l'enclosure parfois opérée par les institutions sur le domaine public et la défense des « communs »²³¹ peut aller jusqu'à conduire, dans le domaine patrimonial, à des actions de sédition à la manière de Wikileaks. L'une des figures emblématiques du mouvement de l'accès ouvert, Aaron Swartz, dont le suicide en 2013 à l'âge de vingt-six ans a profondément ému la communauté mondiale des défenseurs de l'Internet libre, était ainsi poursuivi aux États-Unis pour avoir illégalement mis en ligne des revues téléchargées de la plateforme JSTOR, c'est-à-dire, fondamentalement, des collections numérisées typiques d'une bibliothèque de recherche. En 2008, c'est lors d'une rencontre de bibliothécaires en Italie que s'élabore le *Manifeste pour une guérilla en faveur du libre accès*, d'abord publié sur

229 *Op. cit.*, p.76.

230 B. Barbier, *op. cit.* p. 169-170.

231 Dans un monde numérique, la reproductibilité des contenus pour un coût quasi nul devrait faciliter la circulation de l'information : c'est ce que défend le manifeste de l'association Savoirs Com1 : <http://www.savoirscom1.info/manifeste-savoirscom1/>. Voir aussi l'article de Lionel Maurel, *op. cit.*

le blog d'Aaron Swartz avant d'être retiré²³². « *Il faut nous emparer du savoir, où qu'il soit, effectuer des copies et les partager avec le reste du monde* », écrivait-il pour appeler étudiants, bibliothécaires et scientifiques – bénéficiaires du privilège de l'accès aux ressources numériques de la connaissance – à les partager le plus largement possible. Un « impératif éthique » qui pourrait aussi être celui des bibliothèques face au patrimoine numérique.

232 Cf. Aaron Swartz, *Celui qui pourrait changer le monde : écrits*. Paris : Éditions B42, 2017. L'introduction à ce texte, p. 18, indique que le rôle d'Aaron Swartz dans la conception du manifeste resta un sujet controversé tout au long de la procédure judiciaire.

Conclusion

Depuis les années 1980, un phénomène – à présent bien caractérisé – d’explosion patrimoniale autorise à qualifier de « patrimoine » des objets matériels et immatériels d’une croissante diversité. Le principe de l’élection de leur propre patrimoine par les communautés est désormais reconnu, et l’idée d’un patrimoine numérique, s’inscrivant dans la continuité du patrimoine écrit, est institutionnalisée aussi bien dans l’édifice législatif qui régit le dépôt légal en France que dans les accords internationaux portés par l’Unesco. Pourtant, si l’on en croit l’expérience de la BnF au cours des vingt dernières années, l’acquisition d’une légitimité pour ces nouveaux objets patrimoniaux, qui vont des jeux vidéo sur support aux archives de l’Internet, des ouvrages anciens numérisés aux contenus dématérialisés collectés sur le web, a été une aventure mouvementée.

À chaque étape de ce processus, la question de la continuité a été posée, en particulier par ceux qui ont été, au sein de la BnF, les acteurs de cette patrimonialisation. Continuité de la nature des objets concernés, quand il s’agissait de démontrer que les archives du web permettaient de collecter nombre de contenus publiés auparavant sous forme papier ; continuité des modes d’accès, lorsque Gallica proposait des fonctionnalités de navigation permettant d’appréhender l’objet numérisé dans sa matérialité ; continuité des activités de la bibliothèque, à travers les politiques de conservation élaborées pour SPAR ou le signalement dans le catalogue des ressources électroniques acquises. La mise en évidence d’une filiation entre patrimoine écrit ou documentaire et patrimoine numérique a été un puissant outil de légitimation, lorsqu’il s’est agi de donner une place, d’abord aux documents numériques, puis aux collections numériques dans les missions de la bibliothèque.

Cependant, cette continuité restait indissociable d’une profonde altérité, inhérente à la nature du web et des objets documentaires qu’il génère. Ubiquité, volatilité, déstructuration, massification, internationalisation, transversalité : autant de caractéristiques du web qui ont contraint la BnF à adapter ses processus de travail, ses méthodes de traitement, ses outils et

son organisation pour entrer dans la plus grande révolution depuis Gutenberg, celle qui a conduit à l'émergence d'une culture numérique. Après une période d'expérimentation, jusqu'en 2003, puis une période d'industrialisation et de professionnalisation, jusqu'en 2008, la BnF intègre ensuite pleinement le numérique à ses missions.

Il faudrait étendre cette étude à d'autres établissements, en France et dans le monde, pour mieux comprendre la place de la BnF dans un mouvement que l'on pressent comme général. Une comparaison plus approfondie du processus de patrimonialisation du numérique dans les bibliothèques avec d'autres secteurs comme les archives ou les musées permettrait en outre de mieux saisir la spécificité du patrimoine écrit, et l'héritage de l'appréhension des fonds « anciens, rares et précieux » par les bibliothécaires. On verrait alors sans doute s'esquisser un mouvement de centralisation des gestes professionnels liés au patrimoine numérique, dû à une nécessaire mutualisation des moyens : des dispositifs comme Gallica marque blanche²³³, ou dans les archives des projets comme Vitam²³⁴, apparaissent comme une nécessité lorsque le numérique entraîne à la fois une complexification des techniques documentaires et leur homogénéisation grâce à la convergence des contenus vers un média unique.

Au regard de ces évolutions et des nombreux projets entrepris par la BnF au cours de la période qui a fait l'objet de notre étude, le patrimoine numérique nous apparaît avec deux visages, toujours étroitement liés : d'une part, le numérique comme outil de dissémination, de médiation et de mise en valeur du patrimoine déjà existant, et d'autre part le numérique comme média voyant émerger de nouveaux contenus et objets patrimoniaux. L'étude du processus de patrimonialisation, qu'elle passe par l'observation des outils institutionnels qu'il mobilise, par l'analyse des étapes ou gestes de ce processus, ou par les émotions qu'il suscite, démontre l'impossibilité de démêler ce qui constitue les deux faces d'une même pièce. Parce que la culture numérique irrigue tous les aspects de la société, l'ensemble de ces ressources, qu'elles soient numérisées ou nées numériques, en viennent à faire patrimoine, par l'effet de leur intégration à des collections patrimoniales faisant l'objet d'une gestion adaptée, mais aussi par les usages qu'elles voient émerger.

Derrière les usages de recherche, qui invitent à l'élaboration d'une nouvelle heuristique, d'une critique documentaire de la donnée numérique prenant pleinement acte de la transition

233 Cf. <https://www.bnf.fr/fr/gallica-marque-blanche>.

234 <https://www.programmevitam.fr/>.

des sources vers une nouvelle forme qui engendre ses propres questionnements scientifiques, l'étude – bien trop rapide – des émotions patrimoniales suscitées par le numérique que nous avons proposée dans ces pages suggère une appropriation bien plus vaste. Des amateurs aux ayants droit, des simples curieux aux professionnels, l'appartenance du patrimoine numérique à sa communauté ressort des transports qui s'expriment sur les réseaux sociaux, de la dispute autour de modèles économiques en transition, de la déploration face à une très vive sensation de perte, voire de la sédition lorsque ce nouveau patrimoine est revendiqué dans ses caractéristiques essentielles : appartenant au domaine public, aux « communs » des savoirs, sa diffusion en ligne, ouverte et gratuite, est devenue une obligation et une évidence.

En ce sens, l'émergence d'un patrimoine numérique interroge en profondeur l'identité et les missions de la BnF, pas seulement à l'égard de ces nouvelles collections d'un genre particulier, mais de façon globale dans son rapport à la société et au monde. Dans la vision originelle de Jacques Attali et François Mitterrand, le numérique avait pour mission de réconcilier le patrimoine avec la nation, en utilisant la technologie comme outil de décroisement. Vingt ans d'expérimentations et de pratique ont semble-t-il fini par renverser cette logique, faisant du numérique un nouveau patrimoine dans lequel la nation est appelée à se reconnaître.

De la même manière que cette question a été posée pour d'autres formes de patrimoine, immatériel notamment, on peut s'interroger sur la nature de la relation de la société à son propre présent qui transparaît dans l'élaboration du patrimoine numérique : ce que François Hartog nomme le « régime d'historicité »²³⁵. La prise de conscience de sa grande volatilité, du risque d'une perte massive de mémoire de notre époque, conduit à une entreprise de patrimonialisation qui pose déjà la question de sa masse, de laquelle, au-delà du défi de la préservation, il faudra parvenir à extraire ce qui fait sens pour construire, davantage qu'une mémoire, une véritable histoire, c'est à dire un récit structuré visant à « faire consciemment surgir le passé dans le présent »²³⁶.

Cette constante préoccupation de savoir si le patrimoine numérique que nous constituons aujourd'hui sera suffisant pour permettre la « trouvaille » de demain interroge ainsi notre relation au temps : tout comme les « lieux de mémoire » des années 1980 ou les études

235 François Hartog, *Régimes d'historicité : présentisme et expériences du temps*. Paris : Seuil, 2003.

236 Pierre Nora, « Le retour de l'événement » dans *Faire de l'histoire*, Paris : Gallimard, 1974, cité par F. Hartog, *op. cit.* p. 136.

de la matérialité des sources des années 1990, le patrimoine numérique apparaît comme une construction du présent, où conservateurs et historiens se retrouvent autour d'une conception de l'histoire comme science empirique de l'observation de vestiges et de traces, plutôt que comme connaissance théorique et indirecte du passé²³⁷. On rejoint ainsi l'idée « d'obligation de conservation » pointée par Jean Davallon, la responsabilité de la transmission aux générations futures qui caractérise l'essence du patrimoine. L'élaboration d'un patrimoine numérique, au cours des vingt dernières années, est un élément essentiel de compréhension de la culture numérique, c'est-à-dire de la relation de notre société contemporaine aux pratiques d'inscription matérielle des actions humaines dans des objets porteurs de sens, qu'ils soient textuels ou autres. C'est moins notre rapport au passé qui est en question que la perception du présent et encore davantage, du futur, bouleversée par de nouveaux équilibres géographiques, temporels, sensoriels et intellectuels, encore en construction.

237 Etienne Anheim, dans *La recherche dans les institutions patrimoniales... op. cit.* p. 167.

Bibliographie

- ANGJELI (Anila), BERMÈS (Emmanuelle) et CARON (Bertrand), « Data processing for digital Libraries: the experience of the BnF with Europeana Sounds project », présenté à *Data in libraries: the big picture*. IFLA Satellite meeting. Chicago, University of Chicago Library, 2016. En ligne : <http://library.ifla.org/1961/1/S05-2016-angjeli-en.pdf>.
- BACHIMONT (Bruno), « L'archive du web : une nouvelle herméneutique de la trace ? » *Web Corpora*, 21 juin 2017. En ligne : <https://webcorpora.hypotheses.org/288>. Consulté le 2 septembre 2019.
- BARBIER (Benjamin), *Le processus de patrimonialisation des cultures populaires à l'ère numérique: le cas du jeu vidéo*. Thèse de doctorat, Université Paris VIII, 2016.
- BARBIER (Frédéric), « Patrimoine, production, reproduction », *Bulletin des bibliothèques de France*, n° 5 (2004), p. 11-20.
- BÉQUET (Gaëlle), *Trois bibliothèques européennes face à Google: aux origines de la Bibliothèque numérique (1990-2010)*. Paris : École des chartes, 2014.
- BERMÈS (Emmanuelle), « Text, data and link-mining in digital libraries: looking for the heritage gold », présenté à *Digital Humanities – Opportunities and Risks: Connecting Libraries and Research*. IFLA Satellite Meeting. Berlin, Berlin-Brandenburgische Akademie der Wissenschaften (BBAW), 15 août 2017. En ligne : <https://hal.inria.fr/hal-01643293/document>. Consulté le 26 mai 2018.
- BERMÈS (Emmanuelle), « Éthique et numérisation : les enjeux éthiques de la diffusion numérique », dans *Éthique et patrimoine culturel : regards croisés*, éd. G. CALLEBAUT. Paris : l'Harmattan, 2016.
- BERMÈS (Emmanuelle), *Le Web sémantique en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2013.
- BERMÈS (Emmanuelle), « Bibliothèques, archives et musées : l'enjeu de la convergence des données du patrimoine culturel. », *Documentaliste sciences de l'information*, vol. 48, n° 4 (2012), p. 45-47.
- BERMÈS (Emmanuelle), « Convergence et interopérabilité : l'apport du Web de données / Convergence and interoperability: a linked data perspective », présenté à *IFLA WLIC 2011 : IFLA World Library and Information Congress, 77th IFLA General Conference and Assembly*. Puerto Rico, 2011. En ligne : <https://www.ifla.org/past-wlic/2011/149-bermes-en.pdf>.
- BERMÈS (Emmanuelle), « La gestion des collections numériques : vers une numérisation durable », dans *Manuel de la numérisation*. Paris : Éditions du Cercle de la librairie, 2011.

- BERMÈS (Emmanuelle), « Risk management : methodological principles / Gérer les risques : principes méthodologiques », *International Preservation News (IFLA-PAC)*, n° 41 (juillet 2007), p. 4-8.
- BERMÈS (Emmanuelle), « Les moteurs de recherche : petit précis de mécanique à l'usage des bibliothèques numériques », *Bulletin des bibliothèques de France*, vol. 52, n° 6 (2007), p. 5-11.
- BERMÈS (Emmanuelle), « Des identifiants pérennes pour les ressources numériques : l'expérience de la BnF », *International Preservation News (IFLA-PAC)*, n° 40 (décembre 2006), p. 16-26.
- BERMÈS (Emmanuelle), « Figoblog. Un blog sur Internet, la bibliothéconomie et la confiture de figues ». En ligne : <https://figoblog.org/>. Consulté le 20 octobre 2019.
- BERMÈS (Emmanuelle) et BONICEL (Matthieu), « Biblissima and IIF », présenté à *IFLA WLIC 2016 : IFLA World Library and Information Congress, 82nd IFLA General Conference and Assembly*. Columbus (Ohio), 2016.
- BERMÈS (Emmanuelle), BOULET (Vincent) et LECLAIRE (Céline), « Améliorer l'accès aux données des bibliothèques sur le web : l'exemple de data.bnf.fr », présenté à *IFLA WLIC 2016 : IFLA World Library and Information Congress, 82nd IFLA General Conference and Assembly*. Columbus (Ohio), 2016. En ligne : <http://library.ifla.org/1447/>.
- BERMÈS (Emmanuelle), DUSSERT CARBONE (Isabelle), LEDOUX (Thomas) et al., « Digital preservation at the National Library of France: a technical and organizational overview », présenté à *IFLA WLIC 2008 : IFLA World Library and Information Congress, 74Th IFLA General Conference and Assembly*. Montréal, 2008. En ligne : http://archive.ifla.org/IV/ifla74/papers/084-Bermes_Carbone_Ledoux_Lupovici-en.pdf.
- BERMÈS (Emmanuelle) et FAUDUET (Louise), « The Human Face of Digital Preservation: Organizational and Staff Challenges, and Initiatives at the Bibliothèque nationale de France », *The International Journal of Digital Curation*, vol. 60, n° 1 (2011), p. 226-237.
- BERMÈS (Emmanuelle), FRÉON (Marie-Elise) et MARTIN (Frédéric), « Tous les chemins mènent au numérique : archivage pérenne, numérisation de masse et coopération numérique à la Bibliothèque nationale de France », *Bulletin des bibliothèques de France*, n° 6 (2008).
- BERMÈS (Emmanuelle) et FREYRE (Elisabeth), « Une contribution française à la bibliothèque numérique européenne : Europeana et la Bibliothèque nationale de France », *Digitalia*, vol. 3, n° 1 (2008).
- BERMÈS (Emmanuelle) et ILLIEN (Gildas) « Metrics and Strategies for Web Heritage Management and Preservation », présenté à *IFLA WLIC 2009 : IFLA World Library and Information Congress, 75Th IFLA General Conference and Assembly*. Milan, 2009. En ligne : <https://www.ifla.org/past-wlic/2009/92-bermes-en.pdf>.

- BERMÈS (Emmanuelle) et MARTIN (Frédéric), « Le concept de collection numérique », *Bulletin des bibliothèques de France*, n° 3 (2010), p. 13-17.
- BERMÈS (Emmanuelle) et MOIRAGHI (Eleonora), « Le patrimoine numérique national à l'heure de l'intelligence artificielle. Le programme de recherche Corpus comme espace d'expérimentation pour les humanités numériques », *Revue d'Intelligence Artificielle (RIA)* (à paraître). En ligne : <https://hal-bnf.archives-ouvertes.fr/hal-02122073>. Consulté le 20 octobre 2019.
- BERMÈS (Emmanuelle) et OURY (Clément), « Web 2.0 et mémoire : de la conversation à la conservation », *Documentaliste - Sciences de l'information*, vol. 46, n° 1 (2009), p. 61-63.
- BERMÈS (Emmanuelle) et POUPEAU (Gautier), « Les technologies du web appliquées aux données structurées », dans *Le document numérique à l'heure du web de données. Séminaire Inria, Carnac, 1er-5 octobre 2012*. [Paris] : ADBS éd., 2012. En ligne : <https://hal.inria.fr/hal-00843775/document>, p. 41-84.
- BERMÈS, EMMANUELLE, « Quand le dépôt légal devient numérique : épistémologie d'un nouvel objet patrimonial », *Quaderni : communication, technologies, pouvoir*, n° 98 (2019-2018), p. 73.
- BERTRAND (Paul), « Les digital humanities sont-elles solubles dans le Steampunk ? : Conférence inaugurale », dans *THATCamp Paris 2012 : Non-actes de la non-conférence des humanités numériques*. Paris : Éditions de la Maison des sciences de l'homme, 2012. En ligne : <http://books.openedition.org/editionsmsh/327>. Consulté le 20 octobre 2019.
- BERTRAND (Paul), MAIREY (Aude), GUYOTJEANNIN (Olivier) et al., « L'historien médiéviste et la pratique des textes : les enjeux du tournant numérique », *Actes des congrès de la Société des historiens médiévistes de l'enseignement supérieur public*, vol. 38, n° 1 (2007), p. 273-301.
- BOUCHER (Thierry), *La numérisation des documents imprimés à la Bibliothèque nationale de France*. Mémoire d'études pour le diplôme de conservateur des bibliothèques : Ecole nationale supérieure des sciences de l'information et des bibliothèques, 1994. En ligne : <https://www.enssib.fr/bibliotheque-numerique/documents/62891-la-numerisation-des-documents-imprimees-a-la-bibliotheque-nationale-de-francememoire-d-etude.pdf>. Consulté le 6 octobre 2019.
- CALENGE (Bertrand), *Les bibliothèques et la médiation des connaissances*. Paris : Éditions du Cercle de la librairie, 2015.
- CARDON (Dominique), *Culture numérique*. Paris : SciencesPo les presses, 2019.
- CARDON (Dominique), *À quoi rêvent les algorithmes : nos vies à l'heure des big data*. Paris : Seuil, 2015.
- « Carnet de la recherche à la Bibliothèque nationale de France » En ligne : <https://bnf.hypotheses.org/>. Consulté le 20 octobre 2019.

- CARON (Bertrand), HOUSSAYE (Jordan de la), LEDOUX (Thomas) et al., « Life and Death of an Information Package: Implementing the Lifecycle in a Multi-Purpose Preservation System », présenté à *iPRES 2017 : 14th International Conference on Digital Preservation*. Kyoto (Japon), 2017. En ligne : <https://hal-bnf.archives-ouvertes.fr/hal-01617645>. Consulté le 19 octobre 2019.
- CHANTEPIE (Philippe) et LE DIBERDER (Alain), *Révolution numérique et industries culturelles*. Paris : La Découverte, 2010.
- CLAVERT (Frédéric) et MULLER (Caroline), « Le goût de l'archive à l'ère numérique ». En ligne : <http://www.gout-numerique.net/>. Consulté le 17 octobre 2019.
- COLIN (Nicolas) et VERDIER (Henri), *L'âge de la multitude : entreprendre et gouverner après la révolution numérique*. Paris : Armand Colin, 2015.
- DAVALLON (Jean), *Le don du patrimoine : une approche communicationnelle de la patrimonialisation*. Paris : Hermès science publications : Lavoisier, 2006.
- DEBRAY (Régis), *Les Diagonales du médiologue : Transmission, influence, mobilité*. Paris : Éditions de la Bibliothèque nationale de France, 2014. En ligne : <http://books.openedition.org/editionsbnf/1121>. Consulté le 20 septembre 2019.
- DERAZE (Marianne) et LOYANT (Xavier), « Quand la collection Charles Cros s'expose en ligne. Connaissance et valorisation d'une collection d'appareils de lecture et d'enregistrement unique au monde », *Bulletin de l'AFAS. Sonorités*, n° 42 (septembre 2016). En ligne : <http://journals.openedition.org/afas/2983>. Consulté le 20 septembre 2019.
- DESPRÉS-LONNET (Marie-Pierre), « L'écriture numérique du patrimoine, de l'inventaire à l'exposition », *Culture et musées*, n° 14 (2009), p. 19-38.
- DESRICHARD (Yves), *Cinquante ans de numérique en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2017.
- DOUEIHI (Milad), *Qu'est-ce que le numérique ?* Paris : P.U.F., 2017.
- Émotions patrimoniales*, éd. D. FABRE. Paris : Éditions de la Maison des Sciences de l'Homme, 2013.
- FARGE (Arlette), *Le goût de l'archive*. Paris : Seuil, 1989.
- FARGIER (Nathalie), « Persée, une bibliothèque numérique par et pour les chercheurs », dans *La francesistica italiana à l'ère du numérique*. Gênes : Université de Gênes, SUSLLF (Società Universitaria per gli Studi di Lingua e Letteratura Francese), Institut français d'Italie, 2014. En ligne : <https://hal.archives-ouvertes.fr/hal-01187830>. Consulté le 2 septembre 2019.
- FLICHY (Patrice), *Le sacre de l'amateur : sociologie des passions ordinaires à l'ère numérique*. Paris : Seuil, 2010.

- GATTÉGNO (Jean), *La Bibliothèque de France à mi-parcours : de la TGB à la BN bis ?* Paris : Éditions du Cercle de la librairie, 1992.
- GAUTHIER (Christophe), LAURENT (Natacha), LEVY (Ophir) et al., *Patrimoine et patrimonialisation du cinéma*. Paris : École nationale des chartes, 2020 (à paraître).
- GENIN (Christine), « Le devenir Web de la littérature », *Revue de la BNF*, n° 52, n° 1 (septembre 2016), p. 152-162.
- GILLMOR (Dan), *We the media: grassroots journalism by the people, for the people*. Sebastopol (Canada) : O'Reilly, 2006.
- HAETTIGER (Magali), « Vers la conservation des sites web régionaux », *Bulletin des bibliothèques de France*, n° 4 (2003), p. 77-84.
- HARTOG (François), *Régimes d'historicité : présentisme et expériences du temps*. Paris : Seuil, 2003.
- HEINICH (Nathalie), *La fabrique du patrimoine : de la cathédrale à la petite cuillère*. Paris : Éditions de la Maison des sciences de l'homme, 2009.
- Heritage and digital humanities: how should training practices evolve?* éd. B. SAOUDUFRÈNE. Zürich : Lit, 2014.
- HyperNietzsche : modèle d'un hypertexte savant sur Internet pour la recherche en sciences humaines : Questions philosophiques, problèmes juridiques, outils informatiques*, éd. P. D'IORIO. Paris : Presses universitaires de France, 2000.
- JEANNENEY (Jean-Noël), *Une histoire des médias : des origines à nos jours*. Paris : Seuil, 2011.
- JEANNENEY (Jean-Noël), *Quand Google défie l'Europe : plaidoyer pour un sursaut*. Paris : Mille et une nuits, 2006.
- La Bibliothèque nationale de France : collections, services, publics*, éd. D. RENOULT, J. MELET-SANSON. Paris : Éditions du Cercle de la librairie, 2001.
- La naissance du livre moderne: XIV^e-XVII^e siècles : mise en page et mise en texte du livre français*, éd. H.-J. MARTIN. Paris : Éditions du Cercle de la librairie, 2000.
- La recherche dans les institutions patrimoniales: sources matérielles et ressources numériques*, éd. M. ROUSTAN. Villeurbanne : Presses de l'Enssib, 2016.
- LATRIVE (Florent), *Du bon usage de la piraterie : culture libre, sciences ouvertes*. Paris : La Découverte, 2007.
- LECLAIRE (Céline), « La BnF à l'âge de la multitude », *Bulletin des bibliothèques de France*, n° 13 (janvier 2017), p. 142-143.
- LENIAUD (Jean-Michel), *Les archipels du passé : Le patrimoine et son histoire*. Paris : Fayard, 2002.

- Les collections électroniques, une nouvelle politique documentaire*, éd. P. CARBONE, F. CAVALIER. Paris : Éditions du Cercle de la librairie, 2009.
- Les Lieux de mémoire*, éd. P. NORA. Paris : Gallimard, 1984.
- Les patrimoines en recherche(s) d'avenir*, éd. É. ANHEIM, A.-J. ETTER, G. GLASSON DES-CHAUMES, et al. Nanterre : Presses universitaires de Paris Nanterre, 2019.
- Le tournant patrimonial : mutations contemporaines des métiers du patrimoine*, éd. C. HOT-TIN, C. VOISENAT. Paris : Éditions de la Maison des sciences de l'homme, 2016.
- Le web français de la Grande Guerre : réseaux amateurs et institutionnels*, éd. V. BEAU-DOUIN, P. CHEVALLIER, L. MAUREL. Nanterre : Presses universitaires de Paris Nanterre, 2018.
- Library Linked Data Incubator Group Final Report*. Rapport du groupe W3C Library Linked Data Incubator Group, 2011. En ligne : <https://www.w3.org/2005/Incubator/lld/XGR-llid-20111025/>.
- LOWOOD (Henry), « Memento Mundi: Are Virtual Worlds History? », présenté à *iPRES 2009: the Sixth International Conference on Preservation of Digital Objects*. San Francisco, 5 octobre 2009. En ligne : <https://escholarship.org/uc/item/2gs3p6jx>. Consulté le 20 septembre 2019.
- MACKÉ (Jean-sébastien), « Une édition en ligne : *Le Rêve* de Zola sur Gallica », *Genesis (Manuscrits-Recherche-Invention)*, vol. 23, n° 1 (2004), p. 169-170.
- Manuel du patrimoine en bibliothèque*, éd. R. MOUREN. Paris : Éditions du Cercle de la librairie, 2007.
- MARCETTEAU-PAUL (Agnès), « Le Patrimoine, une valeur d'avenir ? », *Bulletin des bibliothèques de France*, n° 5 (2004), p. 35-38.
- MAUREL (Lionel), « Commun vs. enclosures : quand les bibliothèques publiques verrouillent l'accès aux biens communs informationnels », *BIBLIOTHÈQUE(s)*, vol. 76 (2014), p. 15-21.
- MELOT (Michel), « Qu'est-ce qu'un objet patrimonial ? », *Bulletin des bibliothèques de France*, n° 5 (2004), p. 5-10.
- MOIRAGHI (Eleonora), *Le projet Corpus et ses publics potentiels*. Paris : Bibliothèque nationale de France, 2018. En ligne : <https://hal-bnf.archives-ouvertes.fr/hal-01739730>. Consulté le 20 octobre 2019.
- MOUNIER (Pierre), *Les humanités numériques : une histoire critique*. Paris : Éditions de la Maison des sciences de l'homme, 2018.
- MUSIANI (Francesca), PALOQUE-BERGÈS (Camille), SCHAFFER (Valérie) et al., *Qu'est-ce qu'une archive du web ?* Marseille : OpenEdition, 2019.

- NETZER (Michel), *Les sciences en bibliothèque*. Paris : Éditions du Cercle de la librairie, 2017.
- NOUVELLET (Adrien), BEAUDOUIN (Valérie), D'ALCHÉ-BUC (Florence) et al., *Analyse des traces d'usage de Gallica : Une étude à partir des logs de connexions au site Gallica*. Paris : Télécom ParisTech ; Bibliothèque nationale de France, 2017. En ligne : <https://hal.archives-ouvertes.fr/hal-01709264>. Consulté le 20 octobre 2019.
- Numérisation du patrimoine: quelles médiations ? Quels accès ? Quelles cultures ?* éd. B.N. SAOU-DUFRENE, M. IHADJADENE, D. BRUCKMANN. Paris : Hermann, 2013.
- PALOQUE-BERGES (Camille), « Vers des lieux de mémoire réticulaires ? Construire un patrimoine de la communication des sciences et des techniques du numérique », *RESET. Recherches en sciences sociales sur Internet*, n° 6 (2016). En ligne : <http://journals.openedition.org/reset/839>. Consulté le 17 octobre 2019.
- Patrimoine et passions identitaires*, éd. J. LE GOFF. Paris : Fayard, 1998.
- PÉDAUQUE (Roger T.), *Le document à la lumière du numérique*, éd. J.-M. SALAÜN. Caen : C&F éditions, 2006.
- Pérenniser le document numérique*, éd. L. CALDERAN, B. HIDOINE, J. MILLET. Paris : ADBS éd., 2006.
- POULOT (Dominique), *Une histoire du patrimoine en Occident, XVIII^e-XXI^e siècle : du monument aux valeurs*. Paris : Presses universitaires de France, 2006.
- POUPEAU (Gautier), « L'édition électronique change tout et rien. Dépasser les promesses de l'édition électronique », *Le médiéviste et l'ordinateur*, n° 43 (2004). En ligne : https://archivesic.ccsd.cnrs.fr/sic_00137222. Consulté le 20 octobre 2019.
- POUPEAU (Gautier), « Les petites cases ». En ligne : <http://www.lespetitescases.net/>. Consulté le 20 octobre 2019.
- PRÉMEL (Benjamin), « La numérisation de la presse rétrospective à la BnF », *Le Temps des medias*, vol. 8, n° 1 (décembre 2007), p. 267-278.
- RACINE (Bruno), *Google et le nouveau monde*. Paris : Plon, 2010.
- RACINE (Bruno), *Schéma numérique des bibliothèques* : Ministère de la culture et de la communication, 2010. En ligne : <http://www.ladocumentationfrancaise.fr/rapports-publics/104000143/index.shtml>. Consulté le 2 septembre 2019.
- RENEAR (Allen), MYLONAS (Elli) et DURAND (David), « Refining Our Notion of What Text Really Is. » 6 janvier 1993. En ligne : <http://cds.library.brown.edu/resources/stg/monographs/ohco.html#sec11>. Consulté le 18 mai 2019.
- SALAÜN (Jean-Michel), *Vu, lu, su : les architectes de l'information face à l'oligopole du Web*. Paris : La Découverte, 2012.

- SCHAFFER (Valérie), *En construction : la fabrique française d'Internet et du Web dans les années 1990*. Bry-sur-Marne : INA, 2018.
- SWARTZ (Aaron), *Celui qui pourrait changer le monde: écrits*, éd. L. LESSIG. Paris : Éditions B42, 2017.
- Temps et temporalités du Web*, éd. V. SCHAFFER. Nanterre : Presses universitaires de Paris Nanterre, 2018.
- TRELEANI (Matteo), *Qu'est-ce que le patrimoine numérique ? Une sémiologie de la circulation des archives*. Lormont : Éditions Le Bord de l'eau, 2017.
- Vers de nouveaux catalogues*, éd. E. BERMÈS. Paris : Cercle de la librairie, 2016.
- « Web Corpora | Explorer les archives de l'internet à la BnF ». En ligne : <https://webcorpora.hypotheses.org/>. Consulté le 20 octobre 2019.
- WAGNEUR (Jean-Didier), « Gallica : la bibliothèque électronique de la BnF : quel accès pour les personnes handicapées visuelles ? », dans *Bibliothèques et publics handicapés visuels*. Paris : Éditions de la Bibliothèque publique d'information, 2014 (Paroles en réseau). En ligne : <http://books.openedition.org/bibpompidou/1500>. Consulté le 30 mai 2019.

Table des matières

Remerciements	3
Introduction	5
PREMIÈRE PARTIE : TECHNIQUE ET MATÉRIALITÉ DU DOCUMENT NUMÉRIQUE, L'INVENTION D'UN OUTILLAGE NOUVEAU (1997-2008)	
11	
I. S'approprier les codes du web (1997-2003).....	14
A. Ubiquité : la bibliothèque face aux usages du web.....	15
B. Volatilité : la bibliothèque face à un médium impermanent.....	17
C. Déstructuration : la bibliothèque entre reproduction et réédition.....	19
II. Se transformer politiquement et structurellement : industrialisation et professionnalisation (2003-2008)	26
A. Changement d'échelle : gérer la masse	27
B. Internationalisation : la bibliothèque face à un média global	29
C. Transversalité : quand le web bouscule l'organisation.....	31
DEUXIÈME PARTIE : COLLECTER, CONSERVER, DIFFUSER : L'ÉMERGENCE DES COLLECTIONS NUMÉRIQUES (2008-2018).....	
35	
I. Peut-on parler de collection numérique ? Les premières pierres d'une légitimation (2008-2011)	38
A. Renouer avec la gestion des collections numériques.....	39
B. Retrouver l'unité des collections grâce à la préservation numérique	42
C. Repenser la visibilité de la bibliothèque sur le web	44
II. Vers l'idée de patrimoine numérique (2012-2014)	47
A. Définir de nouveaux objets patrimoniaux	47
B. Faire évoluer métiers et compétences.....	49
C. Construire la présence en ligne de la bibliothèque	50
III. Collections, corpus, données : la bibliothèque face aux nouveaux usages de la recherche (2014-2019)	54
A. Les corpus : une évolution structurelle de la recherche.....	55
B. Les données : de nouvelles formes d'interaction avec les publics	58
C. Convergence et questions épistémologiques	59
TROISIÈME PARTIE : LE PROCESSUS DE PATRIMONIALISATION DU NUMÉRIQUE EN BIBLIOTHÈQUE..	
62	
I. La patrimonialisation institutionnelle du numérique : de la reconnaissance à la légitimité	65
A. Du patrimoine écrit au patrimoine numérique.....	66
B. Le dépôt légal, outil de patrimonialisation institutionnelle	68
C. L'extension progressive du domaine du patrimoine numérique.....	71
II. Le processus de patrimonialisation numérique en questions.....	73

BIBLIOGRAPHIE

A. Une rupture ou une perte ?	74
B. Désigner, classer, décrire, fouiller ?	75
C. Archive, mémoire, trace : quelle obligation de conserver ?	77
D. Transmettre, montrer : peut-on exposer le patrimoine numérique ?.....	78
III. Émotions et patrimoine à la lumière de la culture numérique	82
A. Une culture numérique ? (le transport).....	83
B. Patrimoine numérique, sphère publique et acteurs privés (la dispute)	85
C. Le patrimoine numérique en actions (la déploration).....	88
D. Le patrimoine vu comme un bien commun (la sédition)	89
Conclusion.....	93
Bibliographie	98
Table des matières	106