

## La numérisation des documents

### Introduction

Fin 2004, Google annonce la numérisation à terme de 15 millions d'ouvrages issus des collections de grandes bibliothèques universitaires américaines et anglaises (projet Google Print / Google Books).

Ce projet (étendu depuis au-delà du monde anglo-saxon) tendait à être une innovation déterminante : il visait à rendre les documents disponibles en texte intégral, accessibles à tous gratuitement, à l'échelle mondiale. Le projet suscite de vives réactions des grandes bibliothèques européennes ligüées autour de la BnF (après la prise de position publique du Président de l'époque, Jean-Noël Jeanneney, dès 2005), qui envisagent un « contre-projet » de numérisation mené en coopération. Les bases de la future Bibliothèque numérique européenne (ouverte fin 2008) sont posées.

Depuis, la BnF a mis en œuvre une vaste politique de numérisation, qui s'incarne dans différents programmes<sup>1</sup>.

Dans ce contexte, quels sont donc les enjeux et les méthodes d'un projet de numérisation en bibliothèque ? Comment ont été créées les « bibliothèques numériques » ? Ont-elles révolutionné les modes d'accès à l'information, ou simplement apporté un complément à l'offre documentaire « classique et matérielle » ?

## 1. Les techniques de numérisation

### 1.1. Définition technique. Mode texte et mode image

La numérisation est la codification numérique des intensités lumineuses et de la colorimétrie d'un document : c'est un processus de conversion d'un signal analogique en un signal numérique (signal chiffré binaire), qui s'est développé dans les années 1990 et s'est répandu très rapidement dans les années 2000.

La représentation numérique d'un texte peut se faire selon deux modalités :

---

<sup>1</sup> Pour plus d'informations : <https://www.bnf.fr/fr/la-numerisation-la-bnf> [consulté le 12 juillet 2024].

- Le texte est considéré comme une image, il est représenté sur un mode photographique qui permet de rendre (à l'identique) la forme des caractères. Le fichier numérique est obtenu par numérisation directe du document. On parle alors de « mode image » ou « bitmap »,
  - Chaque caractère du texte à sa représentation unique sous forme numérique.
- Ce type de document en mode caractères est obtenu soit par saisie directe (outils de type traitement de texte), soit par reconnaissance optique de caractères (ROC ou OCR, Optical Character Recognition). On parle alors de « mode texte » (ou « mode alphanumérique »).

La représentation numérique d'une image (estampe, carte, plan, enluminure, etc.) s'effectue nécessairement en mode image. On trouve cependant une reconnaissance automatique pour les manuscrits avec l'HTR (Handwritten Text Recognition) entraînant le logiciel sur un type d'écriture. Il existe également de la reconnaissance de caractères pour les partitions, l'OMR, reconnaissance optique de musique. Bien que de plus en plus performant, l'OMR présente encore un taux d'erreur élevé qui entraîne un temps de correction conséquent.

Le mode image produit donc un « fac-similé électronique » du document d'origine. Les images bitmap sont constituées d'une grille de points appelés pixels. Les images numérisées sont caractérisées par leur résolution : une résolution élevée est l'un des éléments garantissant une reproduction fidèle de l'original, et une impression de qualité et une possibilité de zoom plus grande.

L'acquisition de l'image numérique s'effectue selon deux techniques principales :

- La prise de vue photographique (page par page) à l'aide d'un appareil photographique numérique,
- La numérisation par balayage : des capteurs balayent le document (page par page) et reconstituent l'image pixel par pixel.

Relativement simple à réaliser, le mode image nécessite cependant une description bibliographique du document, ce qui inclut une indexation indispensable pour un accès aisé au document. Le mode image exclut toute recherche à l'intérieur d'un texte. Par ailleurs, il génère des fichiers importants qui occupent de la place sur les supports de stockage.

Le mode texte offre quant à lui la possibilité d'une recherche plein texte. Il est possible d'intervenir directement sur le document électronique. En revanche, le mode texte ne permet pas de conserver la présentation initiale du document original.

Les bibliothèques numériques proposent souvent des accès au mode image et au mode texte, ce qui permet à la fois de voir les images de la numérisation et de faire une recherche plein texte, sur la page affichée ou dans l'ensemble du texte.

En outre, une ressource numérique est définie par un ensemble de métadonnées : une métadonnée est un ensemble structuré d'informations décrivant une ressource. Les métadonnées permettent une description bibliographique approfondie dans un format normalisé (Dublin Core, TEF, LOM, METS). Elles permettent de rattacher le document à l'original, structurent entre eux les différents fichiers qui constituent le document numérique, permettent la gestion des droits d'accès au document, les droits d'impression, de modification, etc.

## 1.2. La chaîne de numérisation

Les différentes étapes de la numérisation dans une bibliothèque sont, schématiquement :

- La *définition du programme de numérisation* : objectifs et contenus, en lien avec la politique documentaire globale,
- L'*acquisition (ou la location) du matériel de numérisation* : scanners, appareils de prise de vue numérique directe, logiciels et stations informatiques,
- Le *pré-traitement* du document,
- La *reconnaissance du contenu* grâce aux logiciels d'OCR pour le mode texte et/ou « acquisition de l'image » (en mode image),
- La *correction des résultats* en mode texte, retouche des fichiers en mode image,
- Le *traitement et l'archivage* des documents numériques,
- L'*exploitation des fichiers* en fonction des objectifs du programme (mise en ligne sur le site web de la bibliothèque, diffusion, commercialisation, etc.). Les avantages et les progrès constants de la technique permettent des applications de plus en plus larges.

Ce schéma concerne les bibliothèques qui ont la volonté et surtout les moyens de mener elles-mêmes leurs programmes.

Il faut aussi retenir une possibilité toute différente pour les bibliothèques, qui consiste à faire appel à la sous-traitance (plutôt que de procéder elles-mêmes aux campagnes de numérisation). Pour la production des documents numériques, les bibliothèques font appel à des prestataires extérieurs spécialisés en numérisation. Dans ce cas, un cahier des charges très précis doit être élaboré avant le choix du prestataire. Un contrôle qualité rigoureux est indispensable en fin de chaîne.

## 2. Les avantages de la numérisation et les domaines d'application liés

### 2.1. Modification et enrichissement de l'offre documentaire

En partie grâce à la numérisation, les bibliothèques peuvent proposer à leurs publics une offre documentaire plus large et plus dynamique, diversifier les ressources qu'elles mettent à disposition, développer de nouveaux modes d'accès aux collections.

La création de « bibliothèques numériques » permet un accès à distance aux collections conservées et numérisées par la bibliothèque et offre la possibilité de toucher de nouveaux publics et de satisfaire de nouveaux usages (consultation à distance des étudiants et des chercheurs, publics empêchés, etc.).

Les bibliothèques universitaires englobent généralement sous l'appellation « bibliothèque numérique » non seulement les fonds qu'elles numérisent mais aussi leurs abonnements électroniques payants, leurs bases de données, etc. Il est donc important de faire la distinction entre ces différentes « briques » d'information.

Dans ce cadre, les fonds numérisés s'intègrent dans les **systèmes d'information** (les SI ou portails) développés par les bibliothèques et SCD. Les exemples pour illustrer ces idées sont désormais très divers.

Aux débuts des bibliothèques numériques, il était facile de citer quelques exemples de bibliothèques novatrices (souvent, les mêmes exemples revenaient, telle la bibliothèque électronique de la BM de Lisieux). Désormais, beaucoup d'établissements proposent une bibliothèque de fonds numérisés, même si l'organisation de ces fonds diffère d'un établissement à l'autre : bases par types de documents (manuscrits, imprimés,

documents graphiques, objets, etc.), présentation de « galeries », bases thématiques (par discipline, par spécialité, etc.).

La nature des fonds numérisés dépend du type de bibliothèque, de sa vocation et de son statut, de sa politique documentaire, des publics qu'elle vise.

Par exemple, pour les BMC (Bibliothèques Municipales Classées), ce sont les fonds patrimoniaux (« anciens, rares et précieux ») qui constituent le socle de la bibliothèque numérique. Les bibliothèques des Grands établissements valorisent les collections anciennes et de recherche qu'elles possèdent. Les SCD (qui possèdent aussi des fonds anciens) incluront dans leur bibliothèque numérique les travaux de recherche de leur université de tutelle, les thèses en ligne, etc. Un CollEx<sup>2</sup> (Collection d'excellence) articulera ses programmes de numérisation avec sa discipline de référence, voire ciblera un pan très spécifique de sa collection.

### Voici ci-dessous quelques exemples significatifs :

- Bibliothèque numérique de la BDIC : <https://argonnaute.parisnanterre.fr/> [consulté le 12 juillet 2024].
- Bibliothèque Diderot de Lyon, CollEx en sciences de l'éducation (cas d'un programme de numérisation ciblé : « Le Journal des instituteurs et des institutrices », collection complète en ligne du périodique pour la période 1858-1940) : <http://numerisation.bibliotheque-diderot.fr> [consulté le 12 juillet 2024].
- Bibliothèque numérique de la Bibliothèque Municipale de Lille : <http://www.bm-lille.fr/patrimoine/presentation-des-collections.aspx> [consulté le 12 juillet 2024].
- Commun patrimoine <https://www.communpatrimoine.fr/communpatrimoine/>, la bibliothèque patrimoniale de plaine commune [consulté le 12 juillet 2024]
- La Bibliothèque numérique aréale, bina, collections patrimoniales numérisées de la BULAC <https://bina.bulac.fr/> [consulté le 12 juillet 2024]

Enfin, deux bibliothèques numériques incontournables :

- Gallica : <http://gallica.bnf.fr> [consulté le 12 juillet 2024].
- Europeana : la bibliothèque numérique européenne : <https://www.europeana.eu/fr> [consulté le 12 juillet 2024]].

---

<sup>2</sup> <https://www.collexpersee.eu/les-projets/> (consulté le 12/07/2024)

## 2.2. Politique de numérisation des patrimoines nationaux et européens

### 2.2.1. Gallica

« Gallica » désigne depuis 1997 l'offre numérique de la Bibliothèque nationale de France et de ses partenaires. C'est une bibliothèque en ligne patrimoniale et encyclopédique, accessible à tous, « *représentative des grands auteurs français et des courants de recherche et de réflexion par-delà les siècles* ».

L'histoire, la littérature, les sciences, la philosophie, le droit, l'économie et la science politique, entre autres, y sont représentés. Gallica offre accès à tous types de supports : imprimés (livres et presse), manuscrits, enregistrements sonores, vidéos, documents iconographiques, partitions, objets, cartes et plans. Depuis 2009, elle propose par ailleurs des collections de partenaires variés (270 en tout). Une partie de ces collections sont directement intégrées dans Gallica. Elle s'est également progressivement ouverte aux bibliothèques numériques de bibliothèques partenaires de la BnF (sélectionnées pour leur complémentarité scientifique et leur compatibilité technique avec les fonds numérisés de la BnF) : un lien est créé avec elles, en moissonnant leurs métadonnées, qui apparaissent dans la liste de résultats, parmi les documents de Gallica. En cliquant dessus, l'utilisateur peut alors consulter le document dans la bibliothèque numérique d'origine.

Aujourd'hui, Gallica est composée de 10 millions de documents (livres, cartes, manuscrits, images, fascicules de presse et revues, partition, enregistrements sonores, etc.).

Pour en savoir davantage sur la constitution de Gallica et son évolution, il est possible de consulter les rubriques d'informations correspondantes sur le site de la BnF :

<https://www.bnf.fr/fr/gallica-et-les-differents-dispositifs> [consulté le 12 juillet 2024].

Gallica n'est cependant pas un cas isolé en Europe et dans le monde. Les grandes bibliothèques nationales offrent souvent un équivalent, plus ou moins vaste :

- Bibliothèque nationale de la diète du Japon : <http://dl.ndl.go.jp> [consulté le 12 juillet 2024].
- « Belgica » : Bibliothèque numérique de la Bibliothèque Royale de Belgique : [http://belgica.kbr.be/fr/accueil\\_fr.html](http://belgica.kbr.be/fr/accueil_fr.html) [consulté le 12 juillet 2024].

Enfin, depuis 2009 existe une autre bibliothèque numérique à forte portée symbolique : la Bibliothèque numérique mondiale de l'UNESCO (World Digital Library) : <https://www.wdl.org/fr> [consulté le 12 juillet 2024]. Elle vise à donner l'accès, gratuitement et au plus grand nombre, « *aux trésors des grandes bibliothèques à l'échelle internationale* ».

### **2.2.2. Europeana**

La BnF s'est positionnée dès 2005 comme un établissement moteur pour la constitution d'une Bibliothèque numérique européenne. Désormais, 3 300 institutions européennes (Institut National de l'Audiovisuel, British Library, Rijksmuseum d'Amsterdam, Le Louvre, etc.) contribuent à l'enrichissement de cette bibliothèque numérique.

Europeana a été inaugurée en novembre 2008 avec le soutien de la Commission européenne. Elle offre l'accès à plus de 58 millions d'objets numériques « *représentatifs du patrimoine historique et culturel de l'Europe* ».

Europeana est le portail numérique du patrimoine européen : il permet de prendre connaissance de l'histoire de l'Europe, de l'Antiquité à nos jours.

### **2.3. Une politique de numérisation soutenue au niveau de l'Enseignement Supérieur et de la Recherche : de la BSN à CollEx-Persée**

De 2009 à 2017, le Ministère de l'Enseignement Supérieur et de la Recherche a piloté la BSN (Bibliothèque Scientifique Numérique) qui était une infrastructure nationale : elle avait pour objectif d'organiser et d'accompagner le déploiement des ressources numériques, au sein des établissements documentaires de l'enseignement supérieur. En mars 2012, un ensemble de recommandations pour la Bibliothèque scientifique numérique (ou BSN) a été validé. Ces recommandations ou actions concernaient de manière générale les ressources numériques et s'organisaient plus précisément autour d'objectifs relatifs à l'acquisition, l'accès et le signalement des ressources numériques. Elle veillait en outre à ce que tout enseignant-chercheur, chercheur et étudiant dispose d'une information scientifique pertinente et d'outils les plus performants possibles.

Elle a évolué en 2018 et devient le Comité pour la science ouverte (CoSO) qui assure la mise en œuvre d'une politique de soutien à l'ouverture des publications et des données de la recherche. (<https://www.ouvrirelascience.fr> [consulté le 12 juillet 2024]).

Les questions de numérisation sont désormais portées par le Groupement d'intérêt scientifique (GIS) CollEx-Persée, créé en 2017, pour une durée de six ans. Il comprend :

- Le portail Persée, créé en 2005 par le ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche,
- Les opérateurs soutenant les politiques nationales en matière d'IST (ABES, CTLES),
- Les structures nationales conservant de grandes collections patrimoniales (BnF, INIST),
- Les Bibliothèques délégataires CollEx
- Les bibliothèques associées CollEx,
- Les collections labellisées.

Chaque année, avec le soutien financier du ministère en charge de l'enseignement supérieur et de la recherche, les établissements membres du GIS décident des projets et évaluent les moyens humains, matériels et/ou financiers qu'ils engagent.<sup>3</sup>

Un des sept programmes est consacré à la numérisation : Collex-Persée s'inscrit dans la lignée de BSN5, avec l'ambition de soutenir une numérisation enrichie et qualitative au service de la recherche. Le groupe de travail « Numérisation » vise ainsi à :

- « Rendre visible l'écosystème de la numérisation, ses principaux acteurs, les communautés de recherche impliquées, les contenus produits et les gammes d'outils existant ;
- Soutenir la mise en œuvre d'une numérisation de qualité articulée à des projets de recherche ou présentant un fort potentiel en terme d'usage scientifique ;
- Soutenir la production de contenus et de données, issues des documents patrimoniaux numérisés, et le développement des outils et des services permettant une utilisation optimale par les chercheurs ;
- Promouvoir la diffusion des contenus de la manière la plus large possible suivant les principes d'une science ouverte ;
- S'assurer de leur visibilité et de leur conservation pérenne. »<sup>4</sup>

Les axes de travail sont ainsi identifiés :

- « Définir et accompagner la mise en œuvre d'une politique de numérisation pour l'enseignement supérieur et la recherche. Le cadre référentiel retenu est celui de l'ouverture des contenus numérisés pour garantir leur visibilité et leur réutilisation ;

---

<sup>3</sup> L'organisation. In. CollEx, Collections d'excellence pour la recherche, disponible en ligne : <https://www.collexpersee.eu/gouvernance/> [consulté le 12 juillet 2024].

<sup>4</sup> Préservation numérique <https://www.collexpersee.eu/groupes/preservation-numerique> [consulté le 12 juillet 2024].

la pérennité par le recours à des normes et des standards ; la collaboration des communautés de chercheurs, des structures patrimoniales et des opérateurs organisés en réseau ; la mutualisation d'outils de production et de diffusion pour assurer l'efficacité du dispositif ;

- Soutenir la numérisation, le traitement, l'enrichissement, la représentation et la diffusion ouverte de documents patrimoniaux et des données qui en sont issues par l'organisation d'appels à projets annuels ;
- Identifier et promouvoir les outils et les méthodes qui contribuent à la production, la diffusion ouverte, l'interopérabilité, l'usage scientifique et la préservation du patrimoine numérisé ;
- Identifier les principaux acteurs qui travaillent sur les contenus et fédérer une communauté ;
- Définir un ensemble de bonnes pratiques sur les plans techniques, documentaires et juridiques et émettre des préconisations en se fondant sur la documentation existant en France comme au niveau international. »<sup>5</sup>

### **Un objectif ciblé de BSN était la valorisation, préoccupation toujours importante.**

La numérisation rend possible la diffusion élargie des fonds dont l'accès était jusqu'alors restreint à certaines catégories de publics :

- Fonds anciens ou précieux numérisés pour la diffusion à destination de nouveaux publics,
- Mise en œuvre d'expositions en ligne pour les collections patrimoniales,
- Mise en vitrine : ciblage d'une partie des fonds, communication externe autour d'un « document inestimable », présentation sous forme numérisée d'un manuscrit précieux nouvellement acquis, etc.
- Mise en valeur de la littérature grise d'une université,
- Mise en valeur du patrimoine scientifique de l'enseignement supérieur et de la recherche.

Un programme de numérisation permet de faire sortir de l'oubli un fonds, de le faire connaître au-delà du cercle restreint des seuls initiés.

La question de la valorisation se porte aujourd'hui également sur la valorisation des usages, innovants ou créatifs, des documents numériques par la communauté des utilisateurs des bibliothèques numériques. De la même manière, la valorisation des

---

<sup>5</sup> *Ibid.*

documents numérisés utilise l'ensemble du panel des outils de communication en bibliothèque, notamment ceux du web participatif.

## **2.4. La numérisation dans la chaîne de conservation**

En matière de conservation, la numérisation véhicule un certain nombre de fausses idées. En particulier, les non avertis peuvent croire que la numérisation permet de gagner de la place dans les magasins : c'est le plus souvent faux ! Une fois numérisé, un document conservé à la bibliothèque n'est le plus souvent pas pilonné. Il reste dans les collections.

Dans la chaîne de conservation des documents, la numérisation intervient comme support de substitution et de valorisation.

Comme tout autre document secondaire, le document numérique permet d'éviter ou de limiter la communication au lecteur du document original. En ce sens, la numérisation participe donc à la préservation des fonds de conservation, puisqu'elle offre une alternative à leur communication directe.

En revanche, en tant que tel, un document numérique peut difficilement être considéré comme un véritable support de conservation : le recul est encore insuffisant en termes de pérennité du support et l'évolution des matériels de lecture trop rapide peut conduire à l'impossibilité de consulter des documents numériques de « première génération » ainsi que l'évolution constante des formats.

En fait, le domaine de la conservation est, typiquement, l'illustration des paradoxes de la numérisation : porteuse d'avenir d'un certain point de vue, la numérisation présente en retour un certain nombre de limites et de contraintes.

## **3. La numérisation : limites et contraintes**

### **3.1. Conservation et archivage pérenne**

La numérisation continue de soulever des interrogations, en particulier du point de vue de l'archivage et de la préservation des documents numériques : ils doivent être enregistrés et stockés en suivant des normes (ISO / AFNOR) qui amélioreront leur longévité, protégeront les données et éviteront les dégradations liées aux transferts de

fichiers. Les fichiers numériques doivent être archivés dans des versions non compressées, et à haute résolution.

La complexité de ces questions a engendré la mise en place d'une politique nationale dédiée à l'archivage numérique pérenne. C'est pourquoi, un segment de la BSN (BSN 6) a été consacré à l'archivage pérenne (voir encadré ci-dessous).

Un des objectifs de BSN – 6 a donc été de déterminer les moyens de rationaliser les pratiques de numérisation en cours, à l'échelle nationale, et de réduire les coûts souvent très élevés de fonctionnement.

### **L'archivage numérique pérenne, de quoi s'agit-il ?**

L'archivage numérique pérenne est un procédé qui vise à conserver les documents informatiques "sur le très long terme", c'est-à-dire plus de 30 ans. Cela ne peut se faire localement au niveau d'une bibliothèque en raison du niveau de technicité requis pour le suivi et des coûts impliqués. Par ailleurs, les données ayant vocation à être conservées plus de 30 ans doivent être sélectionnées suivant des critères précis et partagés : un horizon de 10 ans est actuellement suffisant pour bien des documents dont les originaux sont conservés dans de bonnes conditions, puisque les techniques de numérisation sont aussi appelées à évoluer.

Des institutions telles le CINES (Centre informatique national de l'enseignement supérieur) proposent un service d'archivage pérenne.

L'archivage numérique pérenne a, selon le CINES, 3 objectifs principaux :

- 1- conserver le document numérique
- 2- le rendre accessible
- 3- en préserver l'intelligibilité

Un service d'archivage à long terme recueille les données à conserver. Si l'on veut conserver à long terme la copie numérisée d'un document, on doit se poser plusieurs questions :

- Quelle sera l'utilité du fichier 30 ans plus tard ?
- Comment sera-t-il accessible et exploitable ?
- Quel format choisir pour la conservation ? On trouvera à titre d'exemple une liste des formats actuellement considérés comme durables sur le site du CINES. Cela ne veut pas dire que ces formats sont éternels. Il faudra certainement faire un jour migrer les données dans un format différent.
- Comment décrire le fichier (choix des métadonnées) ?

La Bibliothèque nationale de France s'est dotée d'un outil pour mettre en œuvre ses politiques de préservation numérique, lancé en 2010, SPAR (Système de Préservation et d'Archivage Réparti). *« Au-delà d'un simple stockage sécurisé, le système SPAR cherche à garantir la préservation des documents numériques : à faire en sorte que l'information reste lisible, compréhensible et réutilisable sur le long terme, même si l'environnement technique et humain dans lequel ces documents ont été produits change. Comme il est souvent difficile de réparer des documents numériques après dégradation ou de les relire après disparition des matériels et logiciels d'époque, l'effort porte sur les opérations préventives de veille et de migration. »*<sup>6</sup> La BnF propose une prestation d'archivage numérique, en ouvrant SPAR à d'autres organisations.<sup>7</sup>

## 3.2. Des limites techniques

### 3.2.1. Les outils et matériels utilisés dans la chaîne de numérisation

La qualité technique des scanners, appareils de prise de vue numérique, stations de numérisation, logiciels est déterminante sur le résultat final.

De plus, les logiciels de reconnaissance optique de caractères (OCR, pour le mode texte) ne sont pas encore totalement fiables. Les logiciels OCR haut de gamme peuvent atteindre un taux d'exactitude de 99,99 %, ce qui ne représente pas plus d'une erreur pour 10 000 signes. Les fabricants améliorent régulièrement leurs produits, mais l'OCR reste peu adapté aux documents comportant de nombreux noms propres, des formes grammaticales anciennes ou écrits en plusieurs langues. Le recours aux OCR implique éventuellement un travail de relecture et de correction des fichiers si ceux-ci doivent être lus ou édités.

### 3.2.2. Un environnement informatique adapté

Il est nécessaire avec les portails web / les SI, de disposer d'environnements informatiques performants. Sinon, la numérisation n'aura pas de retombées

---

<sup>6</sup> SPAR (Système de Préservation et d'Archivage réparti). In. Le site de la BnF, disponible en ligne : <https://www.bnf.fr/fr/spar-systeme-de-preservation-et-d-archivage-reparti> [consulté le 14 juillet 2024]. Pour en savoir plus : <https://www.bnf.fr/fr/collections-numeriques-preservees-dans-spar-filieres-et-modeles> [consulté le 12 juillet 2024]

<sup>7</sup> Prestation d'archivage numérique. In. Le site de la BnF, disponible en ligne : <https://www.bnf.fr/fr/prestation-archivage-numerique> [consulté le 12 juillet 2024].

immédiates : quel est l'intérêt pour une bibliothèque de numériser ses collections si elle ne dispose pas d'un portail où elle pourra les signaler, les diffuser, les exploiter ?

Afin de garantir un affichage rapide, la bibliothèque doit là encore se doter d'un environnement informatique des plus récents (réseau à très haut débit) et moduler les formats réservés à la diffusion, à l'archivage, etc.

### 3.2.3. Coûts et moyens humains

Mettre en œuvre des programmes de numérisation représente un investissement financier : l'achat de matériels si le programme est mené par la bibliothèque elle-même ainsi que le coût des contrats ou les marchés de sous-traitance si la bibliothèque fait appel à des prestataires extérieurs.

Les moyens humains : les besoins sont importants en temps de travail et en compétences : relecture des fichiers en mode texte si OCR, travail de retouche des fichiers en mode image, migrations de données et procédures d'archivage, traitement bibliothéconomique des fichiers numérisés, intégration dans la chaîne de diffusion des documents, etc.

## 4. Les paramètres et obstacles juridiques

Une bibliothèque a-t-elle le droit de numériser tout document conservé dans ses collections ? Quelle approche adopter pour concilier la volonté de diffuser tous types de documents et le nécessaire respect du droit ?

Numériser les collections conservées dans une bibliothèque ne relève pas seulement de la maîtrise des techniques de numérisation. La numérisation et l'exploitation des documents produits doivent prendre en compte les droits fixés réglementairement. La possession matérielle par une bibliothèque d'un document dans ses collections ne signifie pas qu'elle possède les droits nécessaires pour en faire ce qu'elle veut.

**Consultez la fiche du Tronc Commun consacrée au droit d'auteur pour avoir l'ensemble des informations nécessaires relatives à cette notion essentielle.**

#### 4.1. Pour numériser, dans la pratique

Les bibliothèques qui mettent en œuvre des programmes de numérisation ont tout naturellement à respecter les textes : soit en ne diffusant pas de reproduction des documents non tombés dans le domaine public, soit en obtenant les autorisations adéquates.

En fait, **c'est l'usage que la bibliothèque fera d'un document numérisé qui déterminera les conditions du droit à appliquer.**

Schématiquement :

La numérisation d'un document **« tombé dans le domaine public »** ne pose aucune difficulté. Lorsqu'un document n'est plus protégé par le droit patrimonial, on peut le numériser et toute utilisation peut être envisagée.

Quand un document **n'est pas dans le domaine public** (mais qu'il reste soumis au droit patrimonial), sa numérisation peut être permise (au titre des « exceptions » mentionnées dans la Loi), mais dans des conditions très encadrées : aux seules fins de conservation, ou pour « préserver les conditions de sa consultation à des fins de recherche ou d'études privées », et strictement dans l'enceinte de l'établissement. Sans autorisation expresse, la bibliothèque ne peut en aucun cas aller plus loin dans l'utilisation du document. En particulier, elle ne peut ajouter le document numérisé dans une bibliothèque numérique consultable en ligne.

Si la bibliothèque souhaite numériser un document encore soumis au droit patrimonial pour le diffuser sur le web (consultation à distance, via une bibliothèque numérique ou via un portail documentaire), elle doit obtenir une autorisation et négocier les droits avec l'auteur ou ses ayants droit. Cela donne lieu à une contractualisation.

Toute cession de droits doit être constatée par écrit. Elle fera l'objet d'un contrat entre la bibliothèque et les ayants droit. Il est important de définir le plus exhaustivement et le plus précisément possible les conditions d'exploitation des œuvres numérisées par la bibliothèque. Chaque usage ainsi que chaque droit nécessaire à cet usage doivent faire l'objet d'une cession explicite. Les bibliothèques veilleront à stipuler l'étendue, la destination, le périmètre et la durée de l'exploitation dans le contrat passé avec les

ayants droit (exemple : consultation sur place et/ou mise en ligne sur un portail, exploitation commerciale, etc.).

En cas de diffusion large, privilégier la numérisation de documents libres de droits reste la meilleure façon, la plus simple et la moins coûteuse de respecter les textes.

#### **4.2. Garantir les droits d'une bibliothèque sur ses fonds numérisés**

Pour les bibliothèques, la question du respect des droits se pose aussi à l'inverse. Il est en effet fréquent que les bibliothèques numérisent et diffusent des documents qui sont leur propriété inaliénable, sur lesquels elles ont leurs propres droits à exercer. Comment peuvent-elles se prémunir d'une utilisation non autorisée de leurs documents ?

En fait, les bibliothèques ne disposent pas pour le moment d'un arsenal suffisant pour protéger leurs collections numériques présentées sur Internet de toute utilisation non autorisée. Il est possible d'inscrire dans les images la mention d'appartenance à la bibliothèque. Cette mesure demeure purement dissuasive.

La numérisation des documents conservés à la bibliothèque offre des possibilités commerciales, comme tout autre support de reproduction : vente aux particuliers, aux chercheurs, aux éditeurs, etc. Mais comment faire lorsque quelqu'un commande la reproduction d'un document pour en faire un usage public (pour l'édition, pour illustrer une campagne publicitaire, pour une émission télévisuelle, etc.) ? Les bibliothèques appliquent parfois une tarification spécifique, et ajoutent au prix de vente de la reproduction une « redevance d'utilisation ». Empêcher une libre réutilisation, commerciale ou non, de documents du domaine public fait débat, avec notamment le concept de *copyfraud* (revendication abusive de droits sur des reproductions fidèles d'œuvres appartenant au domaine public).

### **Conclusion**

La numérisation en bibliothèque n'est plus une nouveauté. En une décennie, un saut à la fois qualitatif et quantitatif s'est opéré dans le domaine. Les bibliothèques numériques patrimoniales (programmes de numérisation de fonds anciens, en BM ou en BU) sont d'ores et déjà bien ancrées. Quant à la politique de numérisation de l'enseignement supérieur et à la question de l'archivage numérique pérenne, elles font partie des objectifs principaux de CollEx-Persée.

Elle participe de la valorisation des résultats de la recherche universitaire, via des bibliothèques numériques adaptées.

Il convient néanmoins de porter un point d'attention sur l'impact environnemental des politiques de numérisation patrimoniale, et sur les enjeux d'écoresponsabilité entourant ce service numérique d'importance dans de nombreuses bibliothèques<sup>8</sup>.

## Bibliographie

RACINE, Bruno. Schéma numérique des bibliothèques.

<https://www.vie-publique.fr/rapport/30997-schema-numerique-des-bibliotheques>

[consulté le 12 juillet 2024].

WEESTEL, Isabelle. « Le patrimoine passe au numérique », BBF, 2009, n°1, pp.28-35,

<https://bbf.enssib.fr/consulter/bbf-2009-01-0028-003> [consulté le 12 juillet 2024].

BERMES, Emmanuelle. FREON, Marie-Elise. MARTIN, Frédéric. « Tous les chemins mènent au numérique », BBF, 2008, n°6, pp.34-39,

<https://bbf.enssib.fr/consulter/bbf-2008-06-0034-006> [consulté le 12 juillet 2024].

LESQUINS, Noémie. TESNIERE, Valérie. « La bibliothèque numérique européenne », BBF, 2006, n°3, pp.68-80, <https://bbf.enssib.fr/consulter/bbf-2006-03-0068-012> [consulté le 14 juillet 2024].

CLAERR, Thierry (dir.), WEESTEL, Isabelle (dir.). Manuel de numérisation. Paris : Editions du Cercle de la Librairie, 2011.

RIVIERE, Pauline, DUBOIS, Olesca, GOBBO, Cécile. NumaHOP, plateforme de gestion des contenus numérisés, Le numérique universitaire des BU, avril 2022, pp.34-35,

[https://www.amue.fr/fileadmin/amue/systeme-information/documents-publications/la-collection-numerique/amue-collection-numerique\\_20.pdf](https://www.amue.fr/fileadmin/amue/systeme-information/documents-publications/la-collection-numerique/amue-collection-numerique_20.pdf) [consulté le 14 juillet 2024].

---

8

<https://bib.vert.es/abf.asso.fr/quelles-perspectives-pour-des-politiques-de-numerisation-patrimoniale-ecoresponsables>  
(consulté le 12/07/2024)