

ÉTUDE PRÉALABLE À UNE OPÉRATION DE NUMÉRISATION

LA DÉMARCHE MENÉE À LA BIBLIOTHÈQUE INTERUNIVERSITAIRE CUJAS

La bibliothèque interuniversitaire Cujas, située sur la montagne Sainte-Geneviève et spécialisée en droit et sciences économiques, a été fondée en 1820. Ses collections, constituées dès cette date, occupent aujourd'hui environ 20 000 mètres linéaires.

La bibliothèque comprend notamment une réserve de livres anciens, fonds unanimement apprécié, car il rassemble en un même lieu les textes fondateurs du droit français et étranger dans les éditions contemporaines des grands juristes. En conséquence, de nombreux lecteurs viennent consulter ce fonds particulier.

Toutefois, la réserve de la bibliothèque Cujas n'obtient pas la reconnaissance qu'elle mérite, du fait de ses difficultés d'ouverture et de consultation qui entraînent une faible utilisation au regard de sa réelle richesse. De plus, les ouvrages qu'elle renferme ont beaucoup souffert,

aussi bien à cause d'inondations exceptionnelles dans les magasins que des dégradations quotidiennes suscitées par les manipulations et les conditions de conservation, longtemps médiocres.

Une mise en valeur de cette réserve s'avérait donc nécessaire. L'idée de recourir à la numérisation est apparue dans ce contexte en 1998; une réflexion méthodologique a été menée dès cette date pour concevoir plus précisément les contours du projet et en déterminer les critères techniques.

Choix des objectifs

La première étape a consisté à fixer les objectifs du projet, c'est-à-dire à essayer de conceptualiser le produit fini et de définir les enjeux qui s'y attachent. Dans le cas de Cujas, il s'agissait tout d'abord de répondre à un souci de conservation des documents par la mise à disposition des lecteurs d'un support de substitution de qualité.

Il fallait ensuite rendre un meilleur service au lecteur pour remédier aux difficultés d'accès aux documents, leur permettre une meilleure appropriation du texte en leur proposant

FRÉDÉRIQUE JOANNIC-SETA

**Conservateur stagiaire
à l'ENSSIB**

* Cet article est extrait d'un mémoire de DCB, soutenu à l'ENSSIB en 1999 par Frédérique JOANNIC-SETA, *Le rôle de la numérisation dans la mise en valeur des fonds patrimoniaux : l'exemple de la bibliothèque interuniversitaire Cujas.*

des facilités de reproduction. Enfin, l'opération de numérisation devait servir de vitrine pour la réserve et, plus généralement, pour l'ensemble de la bibliothèque, afin d'en faire mieux connaître les collections et les activités.

C'est dans ce contexte que la numérisation a été retenue comme solution de mise en valeur idéale, préférable en tout cas à une opération de microfilmage – plus classique –, qui n'aurait pas totalement rempli les attentes attachées à ce projet. Il en résultait donc que la numérisation de Cujas toucherait un nombre relativement important d'ouvrages et serait plutôt destinée à un public d'étudiants ou de chercheurs, trouvant là un nouvel outil d'exploitation des fonds. On pouvait déjà penser, à ce stade, à utiliser les possibilités offertes par la numérisation en matière de recherche textuelle, qui porteraient sur tout ou partie des textes.

Détermination des moyens

Dans un deuxième temps, s'est posée la question des moyens dont on pourrait disposer pour mener à bien cette opération : moyens humains tout d'abord, car il fallait du personnel capable, d'une part, de mener à bien l'étude préalable qui se terminerait par la rédaction du cahier des charges, et, d'autre part, de numériser éventuellement en interne, puis de gérer le produit fini.

Les questions matérielles ont également été évoquées et ont été entre autres résolues par une rapide analyse des infrastructures informatiques. Enfin, la question du financement de cette opération a été soulevée : des crédits accordés dépendraient bien évidemment la taille du corpus final et les solutions techniques retenues.

Choix des documents à numériser

Le postulat consistait en un traitement des ouvrages de la réserve : la détermination de la nature des documents, textuels à 95 %, et des sup-

ports, papier, allait donc de soi. Il a cependant fallu mener une expertise sur l'état physique des livres de la réserve : celle-ci a confirmé le mauvais état des collections, tant sur le plan des reliures, ce qui présentait l'avantage d'avoir en sa possession des ouvrages qui s'ouvrent bien, que sur le plan des papiers, brunis, jaunis et altérés par des taches d'humidité ou de roussi pouvant entraver la qualité de la numérisation.

**PLUSIEURS
CRITÈRES
ONT ÉTÉ RETENUS
POUR LE CHOIX
DES DOCUMENTS,
UNE CERTAINE
ORIGINALITÉ,
LE RESPECT
DE LA
THÉMATIQUE
DÉFINIE
ET UNE UTILISATION
EFFECTIVE
DES OUVRAGES
RETENUS**

Une enquête a été menée parallèlement pour connaître l'état du catalogue des ouvrages de la réserve : il s'est avéré que tous disposaient d'une notice bibliographique et d'une cote ou d'un numéro d'inventaire, mais de qualité inégale. Il devenait donc nécessaire d'harmoniser le catalogue des ouvrages retenus.

Un corpus de plus d'une centaine de titres a été ensuite constitué, c'est-à-dire plus important que le nombre de livres qui seraient effectivement

numérisés, afin de pallier d'éventuels problèmes techniques. Plusieurs critères ont été retenus pour le choix des documents : une certaine originalité des documents, constatée à partir du catalogue rétrospectif de la Bibliothèque nationale de France (BnF) et des listes d'ouvrages juridiques traités dans le cadre de Gallica, le respect de la thématique définie, afin de présenter un panel du droit français ancien dans sa diversité, civil ou canonique, privé ou public, et enfin une utilisation effective des ouvrages retenus par les lecteurs, vérifiée à partir des statistiques de prêt.

Ces différents choix, réalisés au préalable, ont, pour une grande part guidé les choix techniques qui ont été faits dans une seconde étape de la réflexion.

Détermination des critères techniques

Une étude bibliographique a été menée à ce stade pour mieux connaître les dernières évolutions en matière de numérisation, afin de déterminer plus précisément l'orientation à prendre. Elle a consisté en une lecture de documents « classiques », issus de la littérature professionnelle, mais aussi de périodiques français ou étrangers, plus nombreux, portant sur l'informatique en général et le numérique en particulier.

Une veille technologique sur Internet s'est avérée extrêmement profitable, notamment la consultation des sites de la Bibliothèque du Congrès, de la BnF, de la Bibliothèque Laval ou de la Bibliothèque nationale du Canada, qui proposent en ligne de nombreux renseignements sur les critères qu'ils ont retenus en fonction du type de documents à numériser. La connaissance des projets à l'échelon français, sur le site de la Direction du livre et de la lecture en particulier, permettait également d'appréhender le contexte national et les projets en cours¹.

1. Gaëlle BÉQUET, La numérisation des documents patrimoniaux [en ligne]. (Page consultée le 9 novembre 1998). Adresse URL : <http://www.culture.gouv.fr/culture/conservation/>

Cette approche a été complétée par la rencontre de professionnels de la numérisation pour des entretiens théoriques mais aussi pratiques : la visite d'un atelier de numérisation a constitué ainsi une étape particulièrement constructive, dans la mesure où l'on visualise mieux le type de matériel susceptible d'être employé et les managements nécessaires. On peut déjà, à ce niveau, avoir une idée des délais requis pour la numérisation pratique d'ouvrages chez un prestataire.

À l'issue de cette étape, le vocabulaire et les techniques étaient suffisamment maîtrisés pour comprendre les enjeux attachés à chaque mode de numérisation. Un certain nombre de questions se sont posées à ce niveau, pour lesquelles une réponse semblait déjà poindre.

Prospection auprès d'autres bibliothèques

Il semblait alors important pour la bibliothèque Cujas de confronter les expériences de collègues, et de voir quels étaient les prérequis posés par chacun, les réponses qu'ils avaient apportées et éventuellement leurs conclusions sur la pertinence de leurs choix.

L'élaboration d'un questionnaire

Nous avons alors réalisé un questionnaire, organisé suivant une chaîne fictive de numérisation. Quel qu'ait été le type de contact pris avec nos collègues par la suite, c'est toujours ce formulaire qui a été suivi.

Une première rubrique portait sur l'état de l'opération à ce jour et sur des questions plus larges de calendrier, afin de voir l'état d'avancement des travaux et d'avoir une idée plus précise sur les délais nécessaires pour mener à bien ce type de projet. Une deuxième partie réunissait les éléments de présentation du fonds numérisé : état des ouvrages, volumétrie, problèmes de droit éventuels...

Les choix pratiques étaient abordés ensuite, aussi bien pour connaître le

mode de numérisation que pour savoir si on avait fait appel à une société de sous-traitance.

Un quatrième point traitait plus précisément de ce que nous avons appelé « l'organisation intellectuelle des données », c'est-à-dire les accès prévus aux documents numérisés par le biais de notices bibliographiques, d'index ou de tables des matières.

LE QUESTIONNAIRE ÉLABORÉ SE VOULAIT ASSEZ PRÉCIS, POUR POUVOIR FAIRE LE TOUR DE LA NUMÉRISATION APPLIQUÉE DANS LE MONDE DES BIBLIOTHÈQUES FRANÇAISES

Nous avons abordé ensuite les moyens mis à disposition au sein de l'établissement.

La sixième rubrique se proposait, en guise de conclusion, de présenter les résultats du projet – échecs comme réussites.

Ce questionnaire se voulait assez précis, pour pouvoir faire, autant qu'il était possible, le tour de la numérisation appliquée dans le monde des bibliothèques françaises. Il a été utilisé de différentes manières.

L'utilisation du questionnaire

Dans un premier temps, un rendez-vous a été pris avec des membres de l'équipe de numérisation de la BnF. Le succès de Gallica nous semblait

témoigner de la bonne connaissance de cet établissement dans ce domaine et confirmer son rôle d'expertise. Cette rencontre s'est avérée profitable sur plusieurs points : outre la connaissance du projet conçu par la BnF, des questions précises sur nos ambitions attachées à la numérisation des ouvrages de la réserve nous ont permis d'affiner notre vision. Des orientations techniques nous ont aussi été suggérées en fonction de la description que nous faisons de l'état des documents. Enfin, la BnF nous a présenté rapidement les bibliothèques ayant mené des projets similaires aux nôtres et susceptibles de pouvoir témoigner d'une expérience proche.

Le questionnaire a été utilisé ensuite dans le cadre d'entretiens avec des collègues. Cette approche a été fructueuse. Elle a été complétée par l'envoi du formulaire par télécopie à quelques bibliothèques particulièrement actives, comme les bibliothèques municipales de Troyes ou de Lisieux. Les conclusions tirées des réponses ont été profitables : l'accent mis, par exemple, sur l'importance du contrôle de qualité du produit fini par chacun de nos interlocuteurs a été unanimement souligné. Nous avons également rencontré la personne chargée de la numérisation au sein de la Sous-direction des bibliothèques et de la documentation, qui nous a fait part des informations dont elle disposait et nous a encouragés dans les choix évoqués alors².

Sur un plan méthodologique, nous avons souhaité que notre tour d'horizon soit le plus large possible. De chaque expérience, il nous semblait en effet que nous pourrions tirer des éléments de réflexion. Pour toucher les bibliothèques francophones, nous avons donc proposé le questionnaire sur un forum de discussion de bibliothécaires, biblio-fr. Cette démarche

2. Charlette BUREST, À propos de la numérisation, rapport à l'attention de la Sous-direction des bibliothèques et de la documentation, décembre 1998. Ce rapport, mis à jour, pourra être consulté sur le site de la SDEB : www.sup.adc.education.fr/bib/acti/FnNum.htm

n'a pas eu les résultats escomptés, dans la mesure où nous n'avons eu que deux réponses, mais ces réponses étaient tout à fait intéressantes et précises. Cet échec s'explique sans doute par la longueur du questionnaire, qui a pu effrayer. Toutefois, l'initiative a été productive, dans la mesure où elle nous a permis de lier contact avec d'autres conservateurs intéressés par les résultats de cette enquête. Nous nous sommes parfois même rendu compte que ce questionnaire avait été réutilisé par des collègues pour voir s'ils n'avaient rien oublié dans la conception de leur propre projet!

Enfin, cette recherche d'informations auprès d'autres bibliothèques s'est achevée par la participation à la journée organisée sur le thème de la numérisation par l'Association des conservateurs de bibliothèque. Celle-ci ne nous a pas tant été profitable sur le plan des exemples qu'elle présentait, que dans le sens où elle nous a permis à la fois de vérifier la justesse de certaines de nos orientations et de partager l'universalité de certains cas de conscience que nous nous posions, par exemple sur le recours à un microfilmage préalable.

À l'issue de cette démarche prospective, il est apparu que le projet de Cujas de mise en valeur de documents textuels patrimoniaux semblait relativement novateur pour une bibliothèque de cette envergure. Nous avons regretté cependant que l'état d'avancement des projets des bibliothèques contactées ne nous ait pas toujours permis de juger des avantages et des inconvénients des différents modes suivis.

Adaptation des critères techniques aux documents

Suivant en cela les conseils de l'équipe de numérisation de la BnF, la bibliothèque Cujas a décidé de procéder à un test de numérisation afin de voir comment se comportaient ses ouvrages anciens face aux différentes techniques proposées. Nous avons sélectionné huit documents, point trop précieux pour pouvoir les sortir de l'établissement, et

choisis selon les critères imposés par le sous-traitant : format inférieur à A3 et s'ouvrant suffisamment pour une numérisation à plat. Onze pages ont été définies pour faire l'objet de ce test, comme représentatives des ouvrages de la réserve dans leur diversité : formats variant du folio au in-12, papier plus ou moins bruni, plus ou moins marqué de taches d'humidité, présence de notes manuscrites, de gravures, encre plus ou moins effacées...

LES MODES DE NUMÉRISATION RETENUS POUR LE TEST ONT ÉTÉ CHOISIS GRÂCE AU TRÈS PRÉCIEUX DIGITAL FORMATS FOR CONTENTS REPRODUCTIONS

Les modes de numérisation retenus pour ce test ont été choisis grâce au très précieux *Digital Formats for Contents Reproductions*³.

Un double objectif était attaché à ce test : déterminer en premier lieu le mode de numérisation le plus adapté aux ouvrages de Cujas par rapport à leur prix, et ensuite déterminer les ouvrages dont l'état physique ne permettrait pas de subir le type de numérisation choisi et qu'il faudrait écarter du programme.

Ce test a permis de confirmer, si besoin était, qu'il fallait écarter l'idée

de procéder à une numérisation en mode texte par application d'un logiciel de reconnaissance optique des caractères. Il a permis ensuite d'argumenter le choix entre une numérisation en niveau de couleurs, la plus satisfaisante esthétiquement parlant, qui offrait de grandes possibilités en matière de compression, et une numérisation en niveau de gris. Cette dernière solution a été retenue comme offrant le meilleur compromis entre une image de qualité et les impératifs économiques de Cujas. Elle avait aussi l'avantage de pouvoir permettre la numérisation de toutes les pages proposées pendant ce test, et donc, *a fortiori*, de tous les ouvrages de la réserve de Cujas.

Tous ces éléments réunis ont permis de dresser un diagnostic extrêmement détaillé, aussi bien sur les choix techniques qu'intellectuels attachés au projet de numérisation de la bibliothèque Cujas. Grâce à ce diagnostic – qui constitue de fait un pré-cahier des charges –, on a pu présenter un projet structuré à la Sous-direction des bibliothèques et de la documentation.

Pour la bibliothèque Cujas, le travail préalable ne s'arrêtera pas là, même si, à ce stade, les décisions les plus délicates ont été prises. Il restera encore, par exemple, à préparer les documents en vue de leur départ chez un prestataire : comptage du nombre de pages concernées, signalement des anomalies, comme les pages déchirées ou les dépliants, signalement des spécificités de pagination, travail autour des tables des matières qui seront saisies en mode texte...

Toutefois, le projet à cette date est bien lancé et aboutira sans nul doute à la mise sur Internet des pages numérisées que nous vous inviterons alors à consulter. L'ambition de la bibliothèque Cujas sera dès lors de faire vivre cette base de documents numérisés, soit en en augmentant le nombre au cours d'une prochaine campagne, soit en en améliorant l'appropriation par le lecteur en lui proposant de nouvelles fonctionnalités.

3. Carl FLEISCHHAUER, *Digital Formats for Contents Reproductions*, [en ligne]. (Page consultée le 30 septembre 1998). Adresse URL : http://memory.loc.gov/ammem/format_s.html