

Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques

Diplôme de conservateur de bibliothèque

MEMOIRE D'ETUDE

Le rôle de la numérisation dans la mise en valeur des fonds patrimoniaux :
l'exemple de la bibliothèque interuniversitaire Cujas

Frédérique Seta

sous la direction de
T. Lafouge et E. Noël
E.N.S.S.I.B.

1999

Titre :

Le rôle de la numérisation dans la mise en valeur d'un fonds ancien : l'exemple de la bibliothèque interuniversitaire Cujas.

Title :

Digitization for a better exploitation of rare books : the Cujas library's example.

Résumé :

La mise en valeur d'un établissement et d'une partie de ses fonds peut aujourd'hui passer par leur numérisation et leur mise à disposition sur Internet. La question se pose toutefois de la faisabilité de cette mise en valeur sur des collections anciennes, dont les caractéristiques nécessiteront plus particulièrement une réflexion sur le type de numérisation à mener. Le projet de la bibliothèque interuniversitaire Cujas, en fournissant un exemple de mise en valeur des fonds anciens par la numérisation, permettra de présenter le contexte général et de débattre des difficultés de numérisation d'ouvrages de valeur et dégradés.

Abstract :

Nowadays, a library and a part of its collection can be enhanced by digitizing them and by putting them on disposal on Internet. The question whether this improvement is possible in the case of rare books, whose characteristics more particularly require a reflexion on the type of digitization which is to be laid, can however be raised. By providing an example of the improvement of rare books thanks to the digitization, the plan of the Interuniversity Library Cujas is to allow to set out the general context and to debate the difficulties of digitizing valuable and deteriorated books.

Descripteurs :

Bibliothèque Cujas de droit et de sciences économiques (Paris).

Bibliothèques**Fonds spéciaux**Livres rares.

Images, traitement des**Techniques numériques.

*Numérisation. (Descripteur non existant dans Rameau mais utilisé à la bibliothèque de l'ENSSIB).

Keywords :

Libraries**Special collections**Rare books.

Image processing**Digital techniques.

Introduction.

Pour une bibliothèque, disposer de fonds patrimoniaux s'avère souvent un avantage qu'il faut exploiter aussi bien auprès de son public que de ses autorités de tutelle. Témoignage de l'histoire de la constitution des fonds de l'établissement, soit par confiscation, soit par volonté délibérée des acquéreurs de constituer une réserve, ces documents anciens ou précieux assurent un certain prestige à qui les possède et constituent aujourd'hui des outils de travail pour les chercheurs d'autant plus importants que difficiles à trouver, alors que leur intérêt scientifique est reconnu. Revers de la médaille, la possession de documents anciens s'avère la plupart du temps difficile pour la gestion au quotidien d'une bibliothèque. Ils nécessitent en effet des soins particuliers : catalogage spécifique prenant en compte les caractéristiques de l'imprimerie ou de l'édition anciennes, conditions particulières de conservation répondant aux spécificités du papier ou des reliures, conditions particulières de communication exigées par la valeur historique voire économique de ces documents afin d'en éviter le vol ou la dégradation physique. Or, emportées par la nécessité de suivre l'édition récente, de fournir au lecteur les ouvrages les plus demandés, et face à cette lourdeur de gestion des documents anciens, les bibliothèques n'ont pas toujours les moyens humains et financiers, le temps ou les idées qui permettraient de mettre en valeur comme il le faudrait un fonds patrimonial souvent prestigieux, les faisant devenir en conséquence des "bibliothèques au bois dormant"¹. Dans la pratique, cette mise en valeur de documents patrimoniaux peut se traduire de deux façons : d'abord par un effort de préservation de l'existant, se traduisant par une amélioration du conditionnement, par des reliures préventives ou un microfilmage, mais également par une volonté de diffusion voire de publicité, par le biais d'expositions ou par un repérage des fonds anciens dans le catalogue. L'enjeu pour une bibliothèque est de

¹ Expression de M. Melot, reprise par : M.-P. DION. " Une expérience multimédia : le catalogue de la bibliothèque de Valenciennes ", dans *Patrimoine et multimédia : le rôle du conservateur. Actes du colloque organisé par l'École du Patrimoine, tenu à la B.n.F. les 23, 24 et 25 octobre 1996*. Paris, Documentation française-École nationale du Patrimoine, 1997, p. 172.

concilier ces deux aspects indissociables pour la mise en valeur de son fonds ancien, et en conséquence de son établissement.

Pendant longtemps, la reproduction argentique, c'est-à-dire la production de différents supports photographiques, diapositives, microfiches ou microfilms, a été considérée, et l'est encore, comme un moyen alliant préservation et diffusion. Elle offre en effet un support de substitution permettant de ne plus avoir recours à l'original, tout en pouvant être reproduit en de multiples exemplaires à moindre coût et diffusé sans dommages, la durabilité du support, la qualité de restitution en faisant en outre un outil pérenne².

Depuis quelques années, la numérisation apporte une alternative à la reproduction argentique pour les établissements culturels, bibliothèques, archives ou musées, souhaitant allier diffusion et préservation.

Dans les faits, et en dépassant le cadre réduit des seuls documents patrimoniaux, la numérisation paraît largement répandue dans le monde des bibliothèques. Une enquête de la Bibliothèque nationale du Canada sur les ressources électroniques a en particulier signalé l'existence de 33 établissements canadiens sur les 50 réponses, ayant entrepris une numérisation de collections imprimées. Il faut signaler que 64% des réponses faisaient état d'une volonté d'accessibilité accrue et 33% mettaient en avant la fragilité des originaux³. Cette enquête permet de mettre en évidence l'expansion de la numérisation dans un espace géographique donné. Comment expliquer cet engouement rapide qui rappelle le développement accéléré de l'Internet ? Il faut tout d'abord noter que cette nouvelle technique va plus loin que la reproduction argentique : " la constitution de collection numérisées n'est pas la simple reproduction des documents sur un autre support. La qualité des collections réunies doit être fonction de critères de sélection, à la fois scientifiques et fonctionnels : sources de cohérence thématique et multimédia, attention apportée aux finalités des

² J. DEVILLE. (Page consultée le 9 novembre 1998. La reproduction des documents patrimoniaux. Problématique générale, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/conservation/fr/preventi/guide_dll.htm.

usages sur ordinateur, politique de sauvegarde, services nouveaux attendus enfin d'une dématérialisation des œuvres³⁴. De plus, si cette technologie se trouve former un moyen de mise en valeur, c'est aussi de par les multiples aspects qu'elle revêt, qui la mettent au centre de tous les services d'une bibliothèque : " la numérisation ne vient pas se rajouter aux fonctions d'acquisition, de conservation, de communication, de catalogage, de transmission de l'imprimé. [...] La numérisation vient en modifier transversalement chacune des composantes "³⁵.

Plus concrètement, et sans rentrer dans le détail qui sera exposé au cours de ce mémoire, la numérisation présente de multiples avantages : qualité de la restitution, traduction en langage binaire, c'est-à-dire l'impossibilité de dégradation des données, accessibilité et souplesse de la diffusion, ainsi que la possibilité de retravailler le document numérique pour une meilleure appropriation par l'utilisateur. En contrepartie, bien sûr, il faut en signaler les inconvénients : nécessité pour la bibliothèque de se pencher sur cette technologie toujours en cours d'évolution, coûts restant élevés, et surtout problème de pérennité de ces données, dépendant de formats et d'instruments de lecture toujours améliorés et donc toujours modifiés.

A partir de ces quelques constatations, il apparaît toutefois que la numérisation occupe aujourd'hui une place toujours plus importante dans la mise en valeur des fonds et des activités de l'établissement. On peut dès lors se poser deux questions : dans quelle mesure cette technologie ne répond pas à un phénomène de mode, obligeant les bibliothèques à s'équiper de matériel spécifique et coûteux, scanners, écrans de lecture ou logiciels de traitement, à créer des produits électroniques qui ne constituent peut-être qu'un feu de paille ? L'exemple récent du vidéodisque dans lequel le département des estampes de la Bibliothèque nationale de France ou la bibliothèque Sainte-Geneviève ont investi et qui est aujourd'hui

³ S. HAIGH. (Page consultée le 9 novembre 1998). Les ressources numériques dans les bibliothèques canadiennes : analyse d'une enquête de la bibliothèque nationale, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/cidl/fsurvey.htm>.

⁴ Y. MAIGNIEN. " Vers la bibliothèque numérique ", dans *Patrimoine et multimédia : le rôle du conservateur. Op. cit.*, p. 119.

⁵ Y. MAIGNIEN. " La constitution de la collection numérisée de la B.n.F. : vers un nouvel encyclopédisme ", dans *Literary and linguistic computing*, vol. 10, n°1, 1995, p. 51.

abandonné, doit rester en mémoire. Toutefois, même si la numérisation bénéficie incontestablement d'une mode et si les pouvoirs publics poussent vers cette technologie, le vif engouement du grand public pour les nouvelles technologies de l'information et, en particulier pour l'Internet, laisse croire qu'elle sera plus durable et prendra une place sans cesse accrue au sein des bibliothèques.

Deuxième point, qui nous intéresse plus particulièrement ici, peut-on considérer que la numérisation constitue aujourd'hui, et ce en dehors de tout phénomène de mode, un passage obligé pour la mise en valeur, plus spécifiquement, de fonds anciens ? L'exemple du projet de numérisation de la bibliothèque interuniversitaire Cujas nous permettra de nous pencher plus concrètement sur les apports attendus de cette technologie pour la mise en valeur d'ouvrages juridiques anciens, relevant aujourd'hui de l'histoire du droit.

Dans un premier temps, nous essayerons de montrer l'état de la réserve des livres anciens de la bibliothèque Cujas, à la fois de façon statique en en présentant le fonds et dynamique, en le mettant en relation avec son public. La sous-exploitation de cette collection apparaîtra rapidement. Ensuite, pour mieux situer ce projet spécifique, il conviendra de se replacer dans un contexte plus général, à la fois sur un plan technique, mais aussi géographique et institutionnel afin de mieux comprendre les enjeux de ce type d'entreprise. Enfin, nous verrons les choix technologiques et intellectuels à faire dans le cadre d'une numérisation de livres anciens. Nous insisterons particulièrement sur les problèmes de pérennité sous-jacents à ce projet, qu'il s'agisse de la conservation-préservation des fonds précieux ou du support de substitution ainsi créé.

I. La bibliothèque interuniversitaire Cujas.

1. Présentation rapide de la bibliothèque Cujas.

Pour mieux appréhender les enjeux de la numérisation à Cujas, il convient de présenter rapidement le statut de cet établissement pour pouvoir notamment situer celui-ci au sein de son espace géographique et de l'organisation administrative.

Cujas est un organisme interuniversitaire. En découlent deux conséquences. En tant que bibliothèque universitaire, Cujas dépend de la sous-direction des bibliothèques et donc, plus largement du ministère de l'Éducation nationale, de la recherche et de la technologie. Elle profite donc de subventions accordées dans le cadre de la contractualisation par cet organe gouvernemental pour financer des projets ambitieux et transversaux sur une période de quatre ans, avec nécessité de réaliser le programme à la fin du temps imparti. La bibliothèque Cujas a, pour le prochain contrat, actuellement en cours de discussion, proposé différents axes d'amélioration de son fonctionnement et de mise en valeur de ses fonds. La partie qui demandera le financement le plus important reste le chantier de rétroconversion des fichiers manuels des ouvrages entrés avant 1952. Cujas a en outre présenté une démarche de numérisation, toujours dans ce cadre de contractualisation, pour un budget de 312

Kf⁶. Il semble que le projet ait été favorablement reçu, même s'il est encore trop tôt pour savoir s'il sera réellement adopté et dans quelle mesure le financement demandé sera accordé... Remarquons d'ailleurs que l'acceptation d'un programme portant sur la numérisation ou toutes autres entreprises ayant trait aux nouvelles technologies de l'information n'aurait rien de surprenant dans ce cadre, tant le ministre fait de publicité autour de sa volonté de faire rentrer la France de l'Éducation nationale dans le monde des technologies de l'information.

Le deuxième point qu'il faut souligner dans le cas de Cujas est son caractère "interuniversitaire", ce qui implique le rattachement de fait, non pas à une seule université et à son S.C.D. comme le prévoit la loi de 1985, mais les liens avec deux universités, en l'occurrence Paris-I et Paris-II. Le rôle de ces deux établissements apparaît toutefois inégal dans les destinées de Cujas. Paris-I, université de rattachement, reste propriétaire des locaux et donc maître d'ouvrage pour tout ce qui concerne les travaux d'aménagements intérieurs.

Le caractère interuniversitaire de l'établissement se retrouve aussi dans la composition de son public, de haut niveau et regroupant des lecteurs de plusieurs universités parisiennes. Théoriquement, Cujas se veut une bibliothèque de recherche, accueillant les enseignants-chercheurs et les étudiants à partir du troisième cycle. Ses collections sont constituées dans ce sens et les outils mis à disposition révèlent cette ambition. La réserve constitue l'un de ceux-ci, car n'ont réellement besoin des textes des juristes médiévaux ou modernes que les personnes engagées dans des recherches poussées. Dans cette optique, on peut voir la numérisation comme un outil supplémentaire et moderne mis à la disposition des chercheurs. Dans la pratique, Cujas, malgré ses efforts pour limiter son public en fonction de ses collections et de ses ambitions, doit accueillir des étudiants de premier et second cycle, attirés par la proximité de l'établissement par rapport aux lieux de cours⁷ ou parce qu'ils ne

⁶ Cette demande de financement est fondée sur un devis fourni par la société Jouve à la bibliothèque, pour un projet de numérisation en format TIFF, niveau de gris et résolution de 300 dpi, d'environ 100 000 pages, avec saisie en format Marc des notices des ouvrages concernés.

⁷ La bibliothèque universitaire de Paris-I située à Tolbiac apparaît comme excentrée pour les étudiants attachés au quartier latin.

disposent pas de bibliothèque juridique suffisante, comme dans le cas des étudiants de Paris-II. Si elle se refuse à orienter sa politique d'acquisition en direction d'un public qu'elle n'estime pas prioritairement le sien, la bibliothèque doit toutefois consacrer une partie de son budget à l'acquisition de textes de base. C'est autant d'argent qui aurait peut-être été mieux employé, dans le cadre d'une bibliothèque de recherche, à l'enrichissement de la réserve qui apparaît, ainsi que nous allons le voir, comme un parent pauvre.

Enfin, il faut souligner un relatif déclin, depuis quelques années, de Cujas. Pendant longtemps, cet établissement a pu se prévaloir d'une certaine suprématie dans le domaine du droit sur la place parisienne. Le manque de crédits pour l'acquisition d'ouvrages ou la modernisation des installations, le manque de personnel ont fait perdre en partie à Cujas sa prépondérance comme bibliothèque de recherche juridique, aujourd'hui disputée à Paris par la Bibliothèque nationale de France. Si le droit a été longtemps négligé au sein de ce dernier établissement, l'ouverture du site de Tolbiac a été l'occasion de combler les manques dans ce domaine mais aussi de prendre l'initiative en ce qui concerne les nouvelles technologies, avec la numérisation de textes fondateurs du droit ou la mise en place de signets sur les sites juridiques par exemple.

La bibliothèque Cujas est un peu restée en retrait de ces initiatives et souhaite désormais reconquérir le public des chercheurs en droit, à la fois sur le plan national mais plus largement sur le plan européen. Pour la directrice de la bibliothèque, deux sections peuvent servir de vitrine à Cujas, le Cerdoc, centre de recherche documentaire, en développant un rôle de "laboratoire technologique du droit", et la réserve, par la mise à disposition de ses textes originaux sur Internet.

2. Présentation de la réserve de la bibliothèque Cujas

a. La réserve et sa mise en place.

Historique de la formation de la bibliothèque

La bibliothèque Cujas actuelle découle directement des décisions prises pour l'enseignement du droit au début du XIX^e siècle. Si le décret du 22 ventôse an XII, qui réforme en les réorganisant les Écoles de Droit, préconise dans son article 38, § 9 l'établissement de bibliothèques pour faciliter les études des professeurs et des étudiants, celui du 4^e jour complémentaire an XII, art. 71, qui prévoit de ne créer des bibliothèques que dans les “ villes où il n'y aurait pas de grandes bibliothèques ”, remet à plus tard la fondation d'une bibliothèque de droit sur la montagne Sainte-Geneviève.

En 1820, la commission administrative de la Faculté de droit envisage toutefois de créer un local où pourraient travailler étudiants et professeurs sur les livres qui y seraient rassemblés. Il faut attendre cependant 1829 pour qu'enfin quelques étudiants préparant leur thèse aient accès à une véritable bibliothèque, et 1831 pour que tous les étudiants, sans restriction, puissent profiter des ouvrages dans un local spécifique.

On connaît peu de changement avant l'arrivée en 1876 du bibliothécaire Paul Viollet, qui décide de faire de cette bibliothèque encore embryonnaire un outil de travail efficace pour tous les membres de la Faculté. Il fait en sorte d'accroître les fonds : le nombre d'ouvrages passe de 15 000 en 1876 à 42 000 en 1888. Il se lance aussi à la recherche des livres possédés par l'École de droit prérévolutionnaire, mais cette collection, du fait de l'inexistence d'inventaire, reste difficile à appréhender. On peut donc dire que l'essentiel du fonds ancien de la bibliothèque Cujas remonte à cette période de formation, aux efforts de P. Viollet et de ses collègues pour communiquer aux juristes des documents permettant de remonter aux sources du droit, et aux nombreux dons de collectionneurs dont a bénéficié et bénéficie encore cet établissement.

C'est encore à Paul Viollet que l'on doit l'installation de la bibliothèque dans les nouveaux locaux construits en 1897 par l'architecte Noat : les lecteurs disposent désormais d'une vaste salle de 280 places, dont l'entrée donne sur la rue Saint-Jacques.

L'inflation du nombre d'étudiants, passant de 2 604 en 1898-1899 à 11 255 en 1933, la création d'un dépôt de périodiques, qui oblige à diminuer la superficie de la salle de lecture, rend rapidement la situation critique mais il faudra attendre 1958 pour

que la bibliothèque Cujas prenne possession des locaux qui sont encore les siens aujourd'hui⁸.

Constitution de la réserve.

Comme nous l'avons vu, dès le début de la constitution du fonds de la bibliothèque, une partie des efforts avait été dirigée vers la formation d'une collection des textes fondateurs du droit, ce qui impliquait l'achat en antiquariat, prolongé jusqu'aux années 1960, d'ouvrages ne faisant plus depuis longtemps l'objet d'une édition. Toutefois, ces ouvrages semblent avoir été perçus dès le départ comme des outils de travail, et non comme des livres anciens, rares parfois, nécessitant des conditions de conservation spécifique. En réalité, le fait que la réserve soit d'une formation récente traduit bien la conception qu'ont longtemps eu les bibliothécaires de Cujas. Pendant très longtemps les ouvrages étaient placés soit dans un magasin, spécifique, certes, la "grande réserve", mais soumis aux mêmes conditions d'accès que les autres locaux de rangement, soit répartis dans les magasins, parmi les livres les plus communs et les plus récents, à la portée des professeurs et autres personnes habilitées à descendre dans les sous-sols. De plus, la communication n'était soumise à aucune condition particulière et n'importe qui pouvait, sur simple demande, avoir accès aux documents les plus rares et les emprunter. Conséquence prévisible, il manque un certain nombre d'ouvrages, sans qu'une évaluation plus précise puisse en faire l'inventaire exact.

La prise de conscience de la nécessité de préserver ces fonds fragiles découle directement des deux inondations de 1978 et 1981 qui atteignent les magasins. Le bilan s'avère catastrophique : "Deux incunables, de nombreux ouvrages des XVI^e et XVII^e siècles, les plus précieux justement, ont été endommagés au point d'être incommunicables. [...] Les reliures [...], cassées, desséchées, gondolées, éclatées, sont irrécupérables. Si l'on a pu désinfecter les volumes touchés, pas une seule restauration n'a pu être, faute de crédit, envisagée. [...] Même les ouvrages que

⁸ Rambaud J. *Présentation de la bibliothèque Cujas*, document interne, 2 novembre 1973.

l'inondation a épargnés ne sont pas pour autant sauvés de tout danger. L'atmosphère surchauffée et sèche des locaux menace leur intégrité en brûlant les papiers et réduisent les reliures en poussière ”⁹.

Le directeur de la bibliothèque décide alors, en 1981, de transférer d'urgence les fonds anciens dans le local actuel, dit “réserve Pavillon”, dans le bâtiment administratif. Y sont entreposés dans un premier temps les manuscrits, les incunables, les ouvrages des XVI^e et XVII^e siècles, ainsi que les livres modernes rares, c'est-à-dire les ouvrages déjà rassemblés dans la “grande réserve”. Commence aussi à cette occasion un travail de longue haleine consistant à extraire des fonds généraux les documents du XVIII^e et XIX^e siècles. Cette tâche se poursuit encore aujourd'hui, d'abord de façon non systématique, par exemple par l'interception au guichet de demande de communication pour des ouvrages antérieurs à 1880. Récemment, le magasinier spécialisé chargé plus spécialement de la réserve, a aussi découvert une centaine de livres, qu'il a donc remontés à la réserve Pavillon. Enfin, le travail préparatoire à la dernière tranche de rétroconversion du catalogue, qui consiste en un “peignage” systématique des fichiers, devrait permettre de repérer les derniers ouvrages dignes d'entrer à la réserve.

A l'heure actuelle, ce travail de recensement et de réunion des ouvrages dans les deux locaux qui y sont affectés, à savoir la réserve Pavillon et le “Fonds Balogh”, où sont entreposés les folios, s'accompagne d'une restauration progressive : depuis 1993 à mai 1997, 159 volumes ont fait l'objet d'un traitement pour un total de 428 456 F¹⁰.

Il reste beaucoup à faire pour achever la constitution d'une réserve rassemblant des documents conditionnés pour durer. Toutefois, la création de cette section a permis, dans l'esprit du public de distinguer les fonds anciens des fonds généraux, et de pouvoir donner à cette section une réalité concrète par son identification à ce local particulier. On peut donc considérer que cette constitution récente marque le premier effort de mise en valeur du fonds ancien de la bibliothèque Cujas.

⁹ Notes manuscrites pour une réunion à la Direction du Livre le 25 février 1982, p. [2].

¹⁰ Chiffres communiqués par Mlle Mekhtoub, magasinière spécialisée chargée de l'ouverture de la réserve.

b. Intérêt du fonds ancien.

Composition générale du fonds.

Le fonds ancien est, à ce jour, composé d'environ 12 000 ouvrages, dont les deux tiers des XVIII^e et XIX^e siècles. Beaucoup de ceux-ci concernent la pratique du droit, commentaires sur les coutumiers, manuels notariés, ou les œuvres complètes de grands juristes français et étrangers, ... En quelque sorte, “ cet ensemble constitue un outil essentiel pour l'histoire politique, économique et sociale mais aussi pour la recherche juridique la plus actuelle ”¹¹.

Plus particulièrement, on trouve parmi les 10 040 ouvrages recensés en 1996 :

- 4 135 volumes de grand format au Fonds Balogh.
- 5 905 volumes de plus petit format en réserve Pavillon
- 212 dossiers manuscrits, dont le plus ancien, recueil des œuvres de Robert Godefroy sur les Institutes et le dernier titre des Pandectes, date de 1584. Ce sont essentiellement des cours ou notes de cours cédés à la bibliothèque par les professeurs.
- 289 dossiers d'archives rassemblant des documents relatifs aux étudiants ou des tracts politiques communiqués au sein de la Faculté de droit de 1925 à 1943.
- 2 collections complètes du *Droit* et du *Moniteur universel*.

Particularité du fonds imprimé.

L'absence d'outil fiable pour estimer les fonds a rendu ce travail de caractérisation des fonds imprimés assez difficile. Les chiffres communiqués ci-dessous relèvent davantage d'une évaluation globale, qu'il faudra vérifier quand la rétroconversion des fichiers permettra des interrogations plus précises.

Sur le plan des années de parution des ouvrages, l'essentiel du fonds ancien est antérieur au XIX^e siècle, à environ 60 %, si l'on se fonde sur les chiffres de 1996. Les

¹¹ Notes manuscrites pour une réunion de la Direction du Livre, 25 février 1982, p. [1].

ouvrages du XIX^e siècle s'avèrent également nombreux. Pourtant, à y regarder de plus près, une grande partie de ces derniers ne méritent pas une place en réserve : la représentation de cette période dans la collection précieuse de Cujas en est ainsi artificiellement augmentée. Il faut y voir le résultat d'une mise systématique en réserve des ouvrages portant une cote inférieure à 2 000, cote numériquement faible, dénotant en théorie une entrée ancienne des ouvrages dans le fonds de la bibliothèque.

En remontant dans le temps, la réserve comporte 7 incunables, dont le plus ancien date de 1450, et environ 400 ouvrages du XVI^e siècle. On doit donc constater une forte représentation des livres du XVII^e et XVIII^e siècle, ce qui ne nous étonnera pas. En effet, ces deux périodes ont vu se multiplier les éditions juridiques pour faire face à la demande d'un personnel, employé dans les parlements et autres tribunaux d'Ancien Régime, en augmentation croissante, ainsi qu'à la professionnalisation du métier. Le fonds de la réserve apparaît donc comme représentatif de cet aspect précis de l'histoire de l'édition française.

En ce qui concerne les formats, il faut noter la prépondérance des folios, que l'on peut estimer à environ 40%, avec toujours la même limite que ci-dessus, à savoir le mode de composition des fonds rassemblés dans le Fonds Balogh, qui réunit les grands formats dont beaucoup n'ont pas leur place en réserve. Le reste des documents de la réserve se répartit assez équitablement entre les quartos et les octavos, environ 25 % chacun, et quelques plus petits formats : in-12 et in-16, ne constituant pas plus de 10% du fonds. Il faut d'ailleurs noter que les formats indiqués ci-dessus sont les formats bibliographiques classiques et ne correspondent pas toujours à la taille réelle des ouvrages. De plus, il est intéressant de savoir que les ouvrages de la réserve, et surtout les folios, restent en deçà du format A3, considéré par M. Cathaly, de la B.n.F., comme le format le plus grand qui puisse se satisfaire d'une numérisation.

L'état physique des ouvrages s'avère préoccupant, même si le danger de détérioration irrémédiable n'apparaît pas comme imminent. Une étude récente a montré qu'environ 8 à 10% des volumes nécessitent une intervention immédiate, mais

que 25 à 30 % des ouvrages auront besoin d'une restauration importante à moyen terme¹². Plus précisément, et les photocopies présentées en annexe 2 permettront d'en juger, les ouvrages ont subi de plein fouet les effets des inondations qui ont été désastreuses, comme nous l'avons vu, pour les reliures, mais aussi pour le papier. De nombreux livres sont ainsi marqués par des tâches d'humidité d'ampleur variable. Un mauvais conditionnement a également fait vieillir le papier de nombreux documents, ce qui se caractérise par des "roussissures" plus ou moins importantes, parfois sous la forme de taches éparses, parfois sous la forme d'un espace typographique uniformément bruni. Nous en verrons les conséquences sur la numérisation.

Ces documents restent enfin, pour leur très grande majorité, des ouvrages strictement juridiques, même si quelques exceptions confirment la règle, comme une édition des *Contes* de Perrault par exemple. Ce fonds juridique est complété par quelques ouvrages de nature bibliographique, catalogues de livres de collectionneurs, bibliographies juridiques, ou bibliophiliques comme le *Bibliotheca universalis* de Gessner. Il faut également y ajouter quelques ouvrages plus généraux comme une édition de l'*Encyclopédie*. La surreprésentation des ouvrages juridiques, ouvrages professionnels s'il en est, explique le peu d'illustration que l'on retrouve dans ce fonds. Les gravures sur bois ou sur cuivre restent suffisamment exceptionnelles pour qu'on les signale. Les quelques frontispices et les plus nombreux portraits d'auteur constituent en définitive les seules illustrations qui figurent dans ces ouvrages.

En définitive, peut-on parler de richesse quand on pense à la réserve de Cujas ? Dans la majorité des cas, elle rassemble des ouvrages peu précieux de par leur contenu. Une conservatrice de la réserve de la Bibliothèque nationale de France considère d'ailleurs que le prix maximum que peut atteindre un ouvrage de droit s'élève en général à 20 000 F sur le marché bibliophilique. Toutefois, ce prix peut être augmenté de par quelques particularités d'exemplaires : on compte ici des ouvrages portant, par exemple, l'ex-libris de Lamoignon, de belles reliures, dont certaines datées du XVI^e siècle, de plus nombreuses éditions originales. Ces ouvrages ne sont

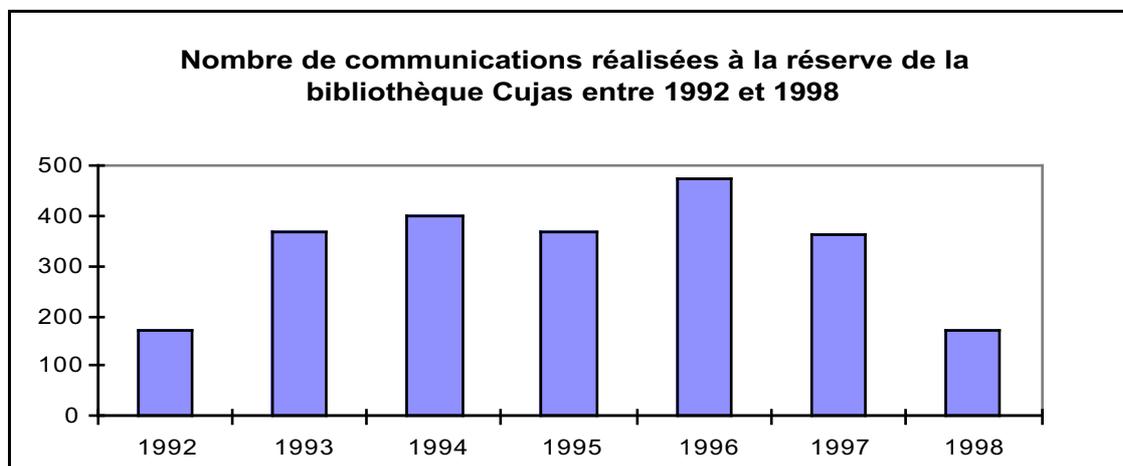
¹² H. LHOUMEAU. *Présentation de la réserve Cujas*, document interne, janvier 1998.

cependant pas très présents dans les fonds de Cujas. Les livres présents restent avant tout les instruments du travail quotidien des robins du XVIII^e siècle, comme en témoignent les nombreuses annotations manuscrites que l'on relève çà et là. L'intérêt du fonds semble surtout de rassembler en un lieu unique un corpus varié d'histoire du droit, civil, canonique, ancien et nouveau, complété par des documents annexes comme des histoires ecclésiastiques. Ne sous-estimons pas toutefois l'originalité des fonds de Cujas par rapport à d'autres bibliothèques. On a pu ainsi estimer récemment qu'environ un quart des ouvrages de la réserve de la bibliothèque Cujas ne se trouve pas à la Bibliothèque nationale de France ; une étude plus systématique, menée sur un échantillon de 106 titres pressentis pour faire l'objet d'une numérisation, a mis en évidence 44 documents ne figurant pas dans le catalogue sur cédérom de la B.n.F. Il faut toutefois un peu pondérer ce dernier résultat : la Bibliothèque nationale de France dispose des textes dans la grande majorité des cas, mais dans une édition différente. Ces chiffres permettent cependant de mettre l'accent sur la richesse de la réserve de Cujas, rassemblant un nombre d'ouvrages qui peut paraître peu important, par exemple par rapport à celle de la bibliothèque Sainte-Geneviève voisine, qui compte 200 000 volumes, mais qui forme un ensemble constitué avec soin.

c. Les usages des lecteurs.

Cette richesse, alliée il est vrai à une proximité de la faculté de droit, fait que la réserve de la bibliothèque Cujas est fréquentée, mais sans doute moins que ne le méritent ses potentialités. Nous reviendrons plus tard sur les modalités de communication des documents de la réserve. Toutefois, pour mieux comprendre le schéma suivant, qui présente le nombre de communications réalisées à la réserve entre 1992 et 1998, il faut savoir que la demande se fait la veille pour le lendemain, par fiche manuelle, dans la salle de lecture générale. Les chiffres qui nous ont permis d'établir ce graphique ont été rassemblés par un vacataire chargé d'évaluer la réserve

en janvier 1998¹³, à partir du “ cahier communication ” tenu manuellement en réserve par le magasinier. Ils sont à comparer avec les chiffres généraux de communication de la bibliothèque Cujas fournis en annexe 7.

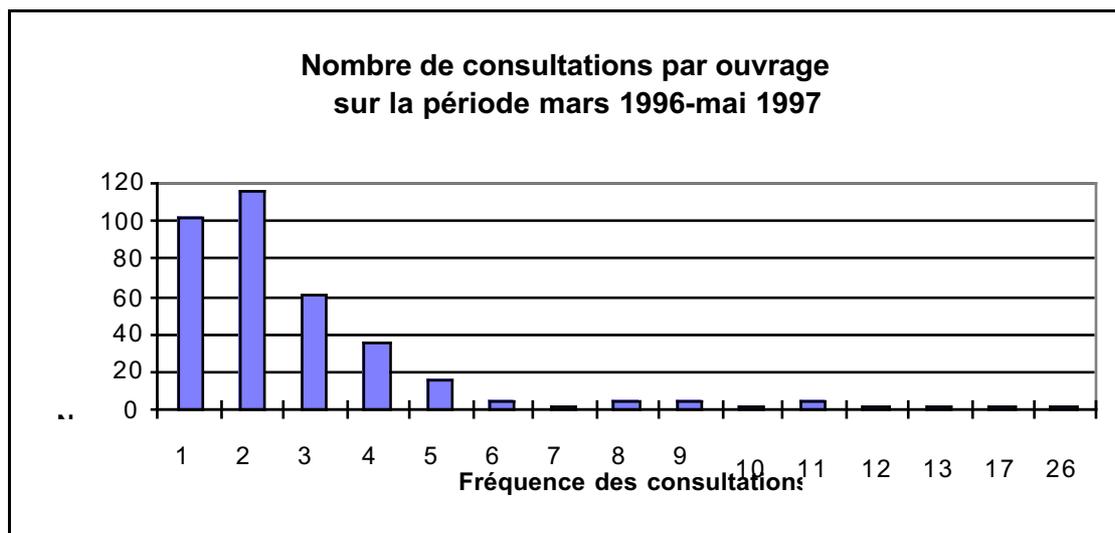


Il faut noter que la diminution du nombre de communications en 1997 correspond essentiellement à des périodes d’ouverture moindres. En effet, la réserve a été fermée pendant l’été 1997. Les chiffres pour 1998 portent uniquement sur la période janvier-mai : on peut estimer qu’en cas d’ouverture normale de la réserve cette année, le nombre de communications aurait été de 408 environ.

Comme on a vu par ce graphique, le nombre de communications annuelles se situe entre 375 et 400, chiffre honorable mais qui rend compte aussi, comme nous le verrons plus bas des difficultés d’accès à la réserve.

Par ailleurs, il est intéressant de noter que les lecteurs demandent souvent les mêmes ouvrages, comme le montre le schéma ci-dessous, établi d’après les mêmes sources statistiques.

¹³ H. LHOUMEAU, *op. cit.*



Le schéma montre que les ouvrages faiblement consultés, de une à deux fois, prédominent. Cependant, un bon tiers des documents sont fortement demandés (c'est-à-dire au dessus de 5 fois, et parfois jusqu'à 26 fois) sur une période d'un an. Ce que l'on remarque cependant, c'est l'importance de ces consultations multiples au sein de la réserve, tendance qu'une étude rapide à partir des registres de consultation verrait confirmée, alors que la très grande majorité des livres de la réserve, près de 96%, ne sont jamais consultés. Cette utilisation d'un faible nombre de documents constitue un terrain favorable pour la création de support de substitution, quel qu'il soit, et *a fortiori* de numérisation, dans la mesure où l'on peut facilement distinguer les ouvrages qu'il convient de préserver du fait de leur forte manipulation. Le choix des livres à traiter, opération souvent délicate de cette entreprise de numérisation, s'en trouve simplifié.

Pour mieux comprendre l'intérêt de la réserve, il faudrait estimer le taux de satisfaction des personnes ayant consulté, ce qui s'avérerait difficile à faire. Le témoignage d'un usager nous autorise toutefois à affirmer l'importance pour eux de ce fonds relativement complet, permettant de comparer facilement plusieurs éditions d'un même ouvrage. Une étude attentive des registres de consultation permettrait

aussi d'établir que la réserve compte des " clients " fidèles, professeurs d'histoire du droit et leurs étudiants en thèse.

Nous avons rencontré un de ces fidèles usagers de la réserve, pour mieux voir si une numérisation des fonds pourrait correspondre à ses besoins. Il apparaît nettement que l'interdiction de photocopie des ouvrages handicape fortement le travail du chercheur, contraint de fait à recopier de larges extraits : le transfert des données sur disquette ou une impression rapide, permis par la numérisation, résoudra pour une grande part ce problème et permettra, de plus, à l'utilisateur d'annoter cette reproduction de texte pour se l'approprier, ce que, semble-t-il, l'utilisateur souffre de ne pas pouvoir faire actuellement.

Le témoignage de cet usager laisse aussi apparaître que les lecteurs viennent irrégulièrement à la réserve dans une année : par exemple, un travail précis nécessitera une forte consultation d'ouvrages sur une période réduite, sur une moitié d'année universitaire, alors que sur l'autre, les obligations professionnelles et l'éloignement géographique empêcheront une consultation sur place. Ce genre de fréquentation par à-coup s'accommode fort mal des horaires réduits d'ouverture pendant la période de presse. Là aussi, la numérisation permettra de résoudre en grande partie ce problème, ainsi que celui de l'éloignement géographique.

d. Les difficultés rencontrées pour la mise en valeur de la réserve.

Même si les lecteurs fréquentant la réserve semblent satisfaits par l'accueil qu'ils y reçoivent et par le contenu du fonds, on doit considérer que la réserve n'est pas encore pleinement mise en valeur actuellement. On peut mettre l'accent sur plusieurs difficultés :

● **Une existence mal connue** : il semblerait tout d'abord que bon nombre des lecteurs de la bibliothèque Cujas ne connaissent pas *a priori* l'existence de la réserve. Il faut dire que le guide du lecteur distribué à chaque nouvelle inscription n'insiste pas vraiment sur cette section particulière qui relève du département de la communication. La réserve n'y est en fait mentionnée que comme un service pratiquant une mise à

disposition différée. Deux lignes indiquent qu'elle abrite des documents anciens mais sans réellement présenter la richesse du fonds. Dans la pratique, la plupart des lecteurs n'apprennent l'existence de la réserve que lorsqu'une demande de communication faite dans la salle de lecture ordinaire leur revient avec la mention "réserve"...

- **Un accès difficile** : du fait du manque de personnel, puisque seule une magasinier spécialisée assure l'ouverture de la réserve tout en s'occupant parallèlement, comme ses collègues, de la communication et la permanence en magasins, la Réserve Pavillon n'est ouverte que du lundi au jeudi de 13h à 17h, soit 16 heures d'ouverture hebdomadaire en moyenne, dont il faut soustraire les périodes de congé du magasinier chargé de l'ouverture. En outre, pour pouvoir consulter les ouvrages de la réserve, il faut les avoir commandés au minimum 24 h. à l'avance en salle de lecture : ce délai est nécessaire pour que le magasinier, qui assure la surveillance de la réserve pendant les heures de consultation, puisse préparer les ouvrages. Certes, quand le lecteur déjà installé en réserve fait la demande d'un ouvrage non réservé mais conservé dans le local Pavillon, le magasinier accède souvent à son désir, mais cette démarche reste exceptionnelle. Dans la pratique, bon nombre de lecteurs demandant en salle de lecture un ouvrage ancien et se voyant renvoyés à la réserve abandonnent là leurs velléités de consultation : on compte ainsi environ 6 abandons par mois sur une moyenne de 40 demandes mensuelles (soit 15%), avec des pics à 21 par exemple pour le mois d'avril 1996.

- **Un manque de lisibilité du contenu de la réserve** : les ouvrages de la réserve sont, pour la majorité d'entre eux, recensés dans le fichier manuel des ouvrages rentrés avant 1952. Or ce fichier, très manipulé, est dans un état assez préoccupant, notamment au niveau des coins déchirés à l'endroit de la cote. Mais surtout, lors de la mise en réserve Pavillon et dans le Fonds Balogh des ouvrages anciens, le report de l'information n'a visiblement pas été fait systématiquement dans ce fichier. Il est difficile pour un lecteur non averti de percevoir ce qu'il va pouvoir consulter sur place ou ce qu'il devra aller voir en Réserve Pavillon. De plus, dans le

cas des ouvrages signalés, la coexistence de plusieurs sigles, retraçant d'une certaine façon l'histoire de la réserve de Cujas, n'améliore pas la lisibilité : " Grande réserve ", " réserve ", " réserve Pavillon " ou " R.P. " ! La conversion rétrospective de ce fichier manuel, à l'ordre du jour à la bibliothèque Cujas, devrait remédier à ces problèmes. Pour les ouvrages anciens entrés après 1952, peu nombreux, l'informatisation a déjà permis une amélioration.

● **Les conditions de consultation** : Elles ne sont réellement pas faites pour attirer le lecteur. Ceux-ci ont à leur disposition six petites tables, dont seulement deux à proximité d'une prise utilisable pour ordinateurs portables. On doit y ajouter une grande table, la seule qui permette réellement de consulter des folios tout en prenant des notes... Cet équipement s'avère insuffisant et il arrive bien souvent à la magasinier de céder le bureau sur lequel elle travaille. En outre, comme le local de consultation sert aussi de magasin, les conditions de lecture apparaissent comme assez exécrables : tables réparties dans l'espace étroit entre deux rayonnages, éclairage artificiel, atmosphère " poussiéreuse " malgré l'entretien attentif. Ces conditions peuvent rebuter des lecteurs non avertis.

En conclusion à cette présentation de la réserve de la bibliothèque Cujas, on peut donc dire que cet établissement possède un fonds ancien tout-à-fait intéressant, qui peut compléter efficacement les collections juridiques de la Bibliothèque nationale de France et qui donne de grandes satisfactions aux historiens du droit et à leurs étudiants. Toutefois, les conditions de consultation, la confidentialité même de la réserve, inconnue ou presque des lecteurs de Cujas, font que cet ensemble apparaît comme peu exploité et peu connu.

II. Le contexte général.

Cette partie veut montrer, à travers une définition de la numérisation et quelques exemples d'entreprises menées dans les bibliothèques, les grandes variétés de projet de numérisation résultant d'objectifs différents.

1. Définition de la numérisation.

A l'heure actuelle, de nombreuses bibliothèques se lancent dans une opération de numérisation¹⁴ d'une partie de leur fonds. Pourtant, à y regarder de plus près, on a l'impression qu'aucune de ces entreprises ne se ressemble vraiment, du fait de divergences d'objectif ou de choix techniques variés. Peut-on alors parler globalement de numérisation ou doit-on considérer qu'il existe plutôt des numérisations ? Ce sera ici notre propos de définir la ou les numérisations.

On pourra tout d'abord essayer de donner une acception globale en comparant deux définitions :

- “ Numérisation : procédé utilisé pour transférer une photo, de la vidéo, du son sur un support informatique. Pour ce faire, ces informations sont transformées en données binaires afin d'être comprises, manipulées et transmises par un ordinateur. La numérisation d'une photo, par exemple, peut se réaliser au moyen d'un scanner, d'une capture vidéo (à partir d'une caméra ou d'un magnétoscope) ou d'une image de synthèse ”¹⁵.

- “ La numérisation fait référence au procédé de traduction en bits d'un document d'information, comme un livre, un enregistrement sonore, une photo ou un vidéo (sic). Les bits sont les unités fondamentales d'information d'un système informatique. Transformer une information en ces chiffres binaires est appelé numérisation. Ce procédé de numérisation peut s'effectuer grâce à diverses technologies existantes ”¹⁶.

¹⁴ Les termes techniques font l'objet d'un glossaire présenté en annexe 1.

¹⁵ L. HERMEL et A.-L. NICOLAS. *Glossaire multimédia et Internet*. Paris, Economica, 1997, p 64.

¹⁶ T. KUNY. (Page consultée le 8 juillet 1998). Introduction aux technologies et aux problèmes de la numérisation, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/pubs/netnotes/fnotes14.htm>.

Ces deux définitions offrent des visions proches de la numérisation en insistant d'ailleurs déjà sur la diversité de celle-ci : diversité dans la mesure où le matériau de départ peut être varié : texte ou enregistrement sonore¹⁷, diversité également au niveau des technologies auxquelles elle fait appel. L. Hermel et T. Kuny insistent d'ailleurs surtout sur la différence de mode de capture des informations et sur l'importance du passage en binaire qui permettra le traitement et la diffusion par l'ordinateur. On peut toutefois dire, pour aller plus loin, que chaque étape de la numérisation offrira une gamme de choix techniques qui font qu'aucune numérisation n'est vraiment semblable à une autre. Plus exactement, " les concepts et les technologies rattachés à la numérisation sont complexes. Il y a un procédé de base qui comporte, à chaque étape, différents jeux de technologies en matériel et en logiciel. La détermination de la technologie adéquate est reliée directement à l'utilisation et au but prévus pour les documents numérisés "¹⁸.

On peut répartir ces choix en suivant les étapes de la chaîne de la numérisation :

- **L'appropriation des données.** Dans le cas de textes, le procédé le plus simple consiste en une saisie directe de l'information sur l'ordinateur. On peut aussi capturer les données grâce à une technique de balayage des objets au moyen d'un scanner et éventuellement procéder à un second balayage, de traitement des données cette fois, par un logiciel de reconnaissance optique de caractères (R.O.C.)¹⁹.

- **Le traitement des données.** Dans le cas des images par exemple, il faudra d'abord choisir l'amplitude des données stockées en mémoire, c'est-à-dire déterminer le nombre de nuances de couleurs à utiliser, mais aussi fixer la résolution de l'image, qui décidera de la qualité, ainsi que du volume, du produit final. Par ailleurs, un texte

¹⁷ A titre d'exemple, la Bibliothèque du Congrès a numérisé 38% de photographies et de films, 6% de cartes géographiques, 5% d'affiches, 18% de manuscrits, 3% de documents sonores, 4% d'images animées et 10% de partitions musicales. G. BEQUET. *La numérisation des collections patrimoniales : les enjeux*. Intervention du 17 juin 1998 au cours du séminaire organisé par l'E.N.S.S.I.B. sur " Conservation et mise en valeur des collections patrimoniales ".

¹⁸ T. KUNY. *Ibid.*

¹⁹ Ou O.C.R. pour les anglo-saxons. Nous ne rentrons pas ici dans les détails, avantages et inconvénients des différentes techniques. Le but ici est d'en proposer un survol rapide pour mieux

peut nécessiter des conversions de signes diacritiques. Dans tous les cas, on peut se trouver dans l'obligation de réaliser des agrandissements ou des réductions, par exemple de diminuer la part des marges pour éviter la conservation de données inutiles. En outre, pour réduire le volume en mémoire d'un fichier, il s'avère souvent intéressant d'utiliser des algorithmes de compression, associés au format de sauvegarde, qui occasionneront ou non des pertes de données.

- **Le stockage.** On peut envisager différents supports de stockage, qui peuvent par ailleurs être distingués des supports d'exploitation : disque dur d'un serveur, bande magnétique, cédérom ou bientôt DVD-rom.

- **La récupération/affichage des données.** Pour une utilisation optimum du produit obtenu, il faudra concevoir des outils de recherche (notices, index) ou de navigation dans le document (table de correspondance, liens hypertexte). On doit aussi se poser la question de la restitution : pour les images, l'affichage sur un écran d'ordinateur demandera une moindre définition qu'une sortie imprimante. De même, une diffusion sur un réseau type Internet exigera des fichiers relativement légers en Kilo-octets.

Il faut ajouter enfin l'importance du respect des normes informatiques, notamment celles de compression, qui seules peuvent garantir une certaine pérennité dans un domaine où la technologie évolue extrêmement rapidement.

Ce rapide survol permet de comprendre les difficultés techniques posées par la numérisation à tout établissement se proposant ce genre d'opération. Il explique aussi pour une grande part que de nombreux projets ne voient pas le jour. Face à cette complexité technique, accentuée par une perpétuelle évolution, il s'avère plus que jamais nécessaire de procéder à une étude préalable poussée afin de créer la meilleure adéquation possible entre les ambitions d'une numérisation et les moyens dont dispose la bibliothèque. Cette volonté d'adapter " l'outil " numérisation à son propre cas explique en grande partie la variété que nous avons signalée plus haut.

comprendre les enjeux dans le cas de la bibliothèque Cujas ou de tout autre établissement se lançant dans ce type d'opération.

2. Les enjeux de la numérisation.

Il existe deux objectifs principaux pour un projet de numérisation :

- **La conservation des documents.** En mettant à la disposition du public des documents numériques, on diminue d'autant la pression sur la collection et on évite la communication des documents originaux. Cela implique toutefois que la qualité des ouvrages et leur état physique soient suffisants pour que le support de substitution puisse remplir son rôle. Une numérisation d'un ouvrage endommagé ne donnera en effet aucun résultat probant. La numérisation dans cette optique joue un véritable rôle de conservation préventive. Comme tout support de substitution, il permet d'assurer une meilleure communication en préservant l'original.

- **La meilleure communication des informations.** Cet objectif doit être compris à la fois dans un contexte local, au sein de l'établissement producteur qui peut offrir à plusieurs lecteurs une consultation simultanée, ou plus large, par le biais de réseau type Internet, ou toutes les possibilités nouvelles qu'offre l'hypertexte. Cette meilleure consultation s'avère lourde de conséquences. Tout d'abord, elle entraîne une meilleure connaissance des collections par les personnes déjà utilisatrices de l'établissement producteur : " parce qu'ils participent des modes de communication actuels [...] et peuvent bénéficier d'un certain " effet de mode ", les projets touchant aux technologies numériques devraient aussi permettre de sauver de l'oubli ou de la disparition nombre de collections [...] " ²⁰. Plus largement, cela permet d'ouvrir les collections à un public plus large, et donc de créer une certaine publicité car " entrer dans une démarche de conception et de production peut [...] faire jouer un rôle actif pour attirer un public plus large vers la culture patrimoniale " ²¹. En poussant l'argumentation jusqu'au bout, on peut même aller jusqu'à parler d'une démocratisation des collections patrimoniales désormais accessibles à tous, sans

²⁰ J.-C. LEYRIS. " Mémoire d'un patrimoine ", dans *Bulletin des bibliothèques de France*, t. 42, n°3, 1997, p. 24.

²¹ L.-M. ALBIGES. " Concevoir et réaliser une banque d'images pour un public non spécialisé ", dans *Du CD-Rom à la numérisation. Développer les documents numériques en bibliothèque*, Villeurbanne, I.F.B., 1997, p. 128.

besoin de justifier d'une recherche ou d'un niveau d'études : ainsi, par exemple, “ a most interesting possibility raised by the digitization of antiquities is teaching university undergraduates the value and pleasure of working with original sources ”²².

A ces deux arguments principaux, on peut rajouter quelques raisons, plus circonstanciées et directement liées à un projet précis : ainsi dans le cas de la Bibliothèque nationale de France, la numérisation doit aussi permettre de développer la lecture et le travail assisté par ordinateur. Notons que cette optique nécessite la mise en œuvre de techniques particulières, numérisation en mode texte en l'occurrence, et ne constitue donc pas, du fait de cette spécificité, un argument recevable pour tout type de numérisation. Il est évident en outre, dans le cas de la Bibliothèque nationale de France, mais aussi, plus largement, dans ceux de tous les projets initiés au plan étatique, que la perspective politique a fortement pesé. De même, le projet sur le patrimoine culturel en région comporte clairement dans la définition de ses objectifs la “ contribution à la modernisation du service public et culturel et [...] la réduction des coûts de mise à disposition de l'information ”, arguments économiques s'il en est.

En conclusion, on se doit d'insister sur le rôle de la numérisation dans la conservation et la consultation des documents, même s'il ne faut jamais perdre de vue les retentissements politiques de tout ce qui touche aujourd'hui aux nouvelles technologies de l'information.

3. Le cadre français : les moyens offerts par l'État.

Comme dans bien d'autres types d'actions, le financement constitue la condition *sine qua non* à la réalisation d'un projet de numérisation. Ce paragraphe a pour objet de donner un rapide historique de la participation de l'État dans ce domaine et d'en définir les cadres administratifs.

²² “ Une plus intéressante possibilité, entraînée par la numérisation des documents anciens, est d'enseigner aux étudiants de 1^{er} cycle la valeur et le plaisir de travailler sur les sources originales ”. H. M. GLADNEY, F. MINTZER, F. SCHIATTARELLA et al. “ Digital access to antiquities ” dans *Communication of the A.C.M.*, vol. 4, n°4, 1998, p. 50.

Pour Jean-Pierre Dalbéra, chef de la mission de la recherche et de la technologie au sein du ministère de la Culture, l'accent mis sur la numérisation au sein de ce ministère découle directement de l'effort de mise en valeur des fonds patrimoniaux par la création de bases de données, par exemple au service de l'inventaire général des monuments et des richesses artistiques de la France. En effet, " au cours de ces vingt dernières années, ces bases de données se sont enrichies et ont été complétées par des banques d'images électroniques permettant d'associer les fiches textuelles aux images des œuvres ou des monuments. Stockées d'abord sur vidéodisques analogiques, ces images ont été progressivement intégrées aux bases de données dès que leur numérisation est devenue possible ". Ainsi, en 1986, le musée d'Orsay a été le premier musée au monde à disposer d'une banque de données et d'images numériques en ligne²³.

On peut dire qu'aujourd'hui, les documents numériques ont acquis une valeur en eux-mêmes et ne sont plus considérés comme de simples illustrations de base de données, au point que " stocker des informations de manière électronique n'est plus un phénomène de mode mais devient une réalité aussi naturelle que l'usage des bases de données " ²⁴. Pour le ministère de la culture, la numérisation se devait de traduire aujourd'hui, aussi bien que les projets autour des bases de données, type Aquarelle²⁵, " le dynamisme des services scientifiques et techniques du ministère et de ses établissements qui s'efforcent en permanence d'explorer de nouvelles voies pour faciliter l'accès à la documentation patrimoniale par le public le plus large, tout en évitant les risques de dégradation des documents originaux causés par de trop fréquentes consultations " ²⁶.

²³ J.-P. DALBERA. (Page consultée le 5 mai 1998). Des bases de données à l'Internet : 25 d'innovations au service de la culture, [en ligne]. Adresse URL : <http://www.culture.gouv.fr/culture/mrt/numerisation/fr/politique.htm>.

²⁴ N. ZOLTAN. *Conception d'un système documentaire avec numérisation de documents pour la documentation régionale de la bibliothèque municipale de la Part-Dieu*. Rapport de stage DESSID 1992 Enssib-Lyon 1, p. 1.

²⁵ Projet européen, fruit d'une collaboration en France entre le ministère de la Culture et l'Institut national de Recherche en Informatique et Automatique (I.N.R.I.A.), cherchant à rendre possible l'interrogation simultanée de plusieurs bases de données réparties, dans la langue de son choix.

²⁶ J.-P. DALBERA, *ibid.*

Plus précisément, et en dehors de toute la réflexion autour de la Bibliothèque nationale de France que nous étudierons plus loin, le ministère de la Culture offre différents types d'aides à la numérisation. Tout d'abord, une ligne de crédit est dégagée sur le budget de la culture pour la numérisation de fonds d'État²⁷, ce qui permet de financer de grands projets à l'échelon national. On peut ainsi citer le projet portant sur la documentation concernant le patrimoine culturel en région. Celui-ci vient d'être lancé dans le cadre du deuxième des six chantiers prioritaires du plan gouvernemental visant à préparer l'entrée de la France dans la société d'information, intitulé " une politique culturelle ambitieuse pour les nouveaux réseaux ". Localement, les Directions régionales d'Action culturelle coordonneront les établissements participant à ce programme²⁸. En 1998, le budget s'élevait à 4 millions de francs²⁹, ce qui permettra de financer 27 programmes³⁰.

Il existe aussi, toujours au sein de ce ministère, une dotation générale de décentralisation qui offre une aide importante aux bibliothèques municipales, car elle peut permettre le financement jusqu'à 40 % du total de la numérisation³¹, soit 22 millions débloqués à condition de respecter certaines directives de la Direction du Livre et de la Lecture (D.L.L.).

Le ministère de la Culture et de la Communication, dont le double objectif, intéressant directement la numérisation, est compris dans son titre, n'est pas le seul acteur politique pour la numérisation. Le ministère de l'Éducation nationale, de la recherche et de la technologie constitue également un partenaire de choix, notamment en offrant aux bibliothèques universitaires des subventions dans le cadre de la contractualisation. Plus spécifiquement, au sein de la sous-direction des bibliothèques, un conservateur est chargé exclusivement de ce dossier. Toutefois, il

²⁷ G. BEQUET. *ibid.*

²⁸ France. Ministère de la Culture et de la Communication.. (Page consultée le 5 Mai 1998). Appel à projets pour la numérisation de la documentation sur le patrimoine culturel en région, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/mrt/numerisation/fr/numerisation_programmes.htm.

²⁹ France. Ministère de la Culture et de la Communication,, *ibid.*

³⁰ J.-P. DALBERA. (Page consultée le 9 novembre 1998). Patrimoine culturel et société de l'information, [en ligne]. Adresse URL : <http://www.culture.gouv.fr/culture/mrt/numerisation/fr/intro-generale.htm>.

n'existe pas encore de réelle politique de la numérisation au sein de ce service, comme il peut notamment en exister à la D.L.L. Le rôle de ce conservateur consiste surtout aujourd'hui en la rédaction d'une synthèse des entreprises de numérisation réalisées dans le cadre des établissements relevant de la sous-direction. Il espère que ce rapport servira à l'avenir de modèle et donnera les moyens de projets plus ambitieux et initiés par la sous-direction³².

Rien n'interdit non plus une collaboration entre les deux ministères, comme par exemple autour du projet mené par l'Institut de Recherche et d'Histoire des Textes (I.R.H.T.) et la D.L.L. Cette entreprise vise à microfilmer et numériser systématiquement toutes les enluminures des manuscrits médiévaux aussi bien dans le cadre de bibliothèques municipales comme celle de Lyon ou de bibliothèques universitaires comme Sainte-Geneviève³³.

Plus largement, de nombreux projets sont menés à l'échelon européen, notamment grâce à la Direction Générale XIII, qui a par exemple financé pour moitié le projet EDIL. Celui-ci devait permettre la diffusion de documents numérisés, essentiellement périodiques, entre différents établissements européens, avec l'implantation de relais nationaux qui assuraient à la fois une fonction technique d'interopérabilité des systèmes et une fonction de système d'information³⁴.

Enfin, on peut citer également une volonté mondiale de mise en commun des ressources, dont Internet n'est qu'une illustration empirique, par exemple avec le projet "Bibliotheca universalis", dirigé conjointement par la France et le Japon en 1995 pour définir les "lignes directrices d'un système d'information multimédia

³¹ G. BEQUET, *ibid.*

³² Entretien avec Mme Buresi, jeudi 29 octobre 1998. Le rapport évoqué ci-dessus a été publié en décembre 1998 sous le titre "A propos de la numérisation".

³³ D. COCQ. *Le rôle des institutions en matière de conservation et de mise en valeur du patrimoine écrit : la politique de la Direction du livre et de la lecture hier et aujourd'hui*. Intervention du 16 juin 1998 au cours du séminaire organisé par l'E.N.S.S.I.B. sur "Conservation et mise en valeur des collections patrimoniales".

³⁴ G. BEQUET. "Bibliotheca universalis. Vers un catalogue commun des documents numérisés accessibles par les réseaux télématiques", dans *Bulletin des Bibliothèques de France*, t. 40, n°5, p. 55.

distribué, donnant un accès virtuel au contenu ”, c’est-à-dire la mise en commun et le classement des différents programmes de numérisation locaux ou nationaux³⁵.

L’ampleur et les ambitions de ces projets nationaux ou supranationaux ne doivent pas nous tromper et ils ne constituent que l’arbre qui cache la forêt. Pour ces programmes bien ciblés, répondant à des objectifs parfaitement établis et avec des obligations techniques prédéfinies, combien de plus petits projets, réalisés avec les moyens propres des établissements ? On peut citer dans ce cas les efforts de O. Bogros à la Bibliothèque municipale de Lisieux, appréciés notamment au Canada par G. Teasdale³⁶, mais la saisie manuelle de textes régionaux par une secrétaire reste un projet local, auquel il manquerait peut-être une coordination et une collaboration plus importante pour en faire un projet d’envergure. De même, la Bibliothèque interuniversitaire de Médecine mène actuellement une politique de numérisation de son fichier manuel ancien mais sur ses fonds propres et sans dépasser le seuil des 300 000 F.

En définitive, si l’aide de l’État est réelle, s’il existe des cadres juridiques, financiers et administratifs, de nombreux projets sont menés indépendamment, preuve d’esprit d’initiative et de créativité, certes, mais avec le risque de manque d’interopérabilité entre les différents programmes.

4. Présentation de quelques exemples de numérisation en bibliothèque.

Afin de présenter les différents modes de numérisation qui peuvent coexister aujourd’hui, nous avons choisi de nous attarder sur trois exemples significatifs, représentatifs chacun d’une conception originale en raison de ses objectifs et donc des moyens mis en œuvre. On peut en effet distinguer trois grands types de numérisation : une opération visant à préserver des documents précieux et en danger, une entreprise visant à valoriser un thème particulier, comme l’histoire d’un pays, et une collection numérisée à but encyclopédique.

³⁵ G. BEQUET. *Op. cit.*, p. 54.

³⁶ G. TEASDALE. (Page consultée le 30 septembre 1998). Le projet Gutenberg, [en ligne]. Adresse URL : http://www.bibl.ulaval.ca/info/pagepers/teasdale/LBQ/lettre_12.htm.

a. La bibliothèque nationale de Tchéquie : la numérisation de documents précieux.

Les projets de numérisation de la Bibliothèque nationale de Tchéquie doivent tout d'abord se replacer dans le contexte du programme de l'Unesco "Memory of the world" et ce depuis 1992, date à laquelle cet établissement a pris en charge le projet "memoriae mundi series bohemia". L'objectif en apparaissait clairement défini : la préservation de documents précieux : "our purpose was clearly defined as manufacturing of replacement media there, where the originals were in bad condition [...] because [...] the damage caused by man was defined as the most dangerous deterioration factor having hard impact on our historical holdings"³⁷. Dans cette optique, deux problèmes se sont rapidement posés : la qualité des images fournies et l'interface utilisateur.

En ce qui concerne la qualité de l'image, il a été fait le choix d'une résolution très élevée, au delà de la perception humaine, permettant une compression avec perte, type JPEG, sans implication sur la consultation du document. Quant à l'interface, il est apparu rapidement qu'elle ne pouvait être conçue qu'en collaboration avec les spécialistes travaillant sur ce type de documents historiques.

Si une première opération de ce type connut le succès en 1995, le développement de l'Internet, en proposant de nouveaux outils, entraînait également une nouvelle réflexion permettant de résoudre à la fois le problème des accès aux différents types de documents, posés par le programme de l'Unesco, et celui de l'archivage des données numériques. Il fut alors fait le choix de s'appuyer sur les possibilités offertes par l'H.T.M.L., avec l'idée de distinguer la description de l'original de sa copie numérique. Ainsi, "the H.T.M.L. is supporting the whole book structure going down from the upper book level to the lower levels as defined by the

³⁷ "Notre objectif était clairement défini comme la création d'un support de substitution là où les originaux étaient en de mauvaises conditions. Les dommages causés par l'homme étaient en effet considérés comme le facteur le plus dangereux de détérioration s'étant répercuté sur nos fonds historiques". A. KNOLL. (Page consultée le 9 juillet 1998). Digitization of old manuscripts in the

describer of the digital copy. [...] The structuring metadata are tagged. In order not to introduce new tags, the H.T.M.L. comment tags are used for this purpose ”³⁸. La bibliothèque espère que le choix de ce langage pour la structuration des données, du fait de sa forte implantation dans le monde, garantira la conservation des données numériques, et de fait la préservation des originaux : “ Our opinion is that the Internet basic language is that tool which can prolong the life of the access to digital data, because access and preservation are in this situation very interdependent issues forming a system the solution of which must be developed in harmony of both ”³⁹. Cette structuration H.T.M.L. est en cours d’application sur un projet initié sur deux ans à compter de début 1997, avec pour but de réécrire les règles de structuration des documents numérisés dans le cadre du projet “ Memory of the world ”, afin de les simplifier et de les rendre applicable à tous types de documents.

Au travers de cet exemple, se pose pour les documents rares et précieux toute la problématique de la conservation des données numériques garantissant de par leur existence l’inutilité du recours au document original et *de facto* la préservation de celui-ci. Le recours à l’H.T.M.L., solution choisie ici, ne peut évidemment être considéré comme une solution miracle.

Dans le cas exposé ci-dessous, les enjeux apparaissent comme différents.

b. La Bibliothèque nationale du Canada : le choix de retracer l’histoire d’un pays.

Résultant d’un partenariat entre l’Institut canadien de microreproductions historiques, la Bibliothèque nationale du Canada, la Bibliothèque universitaire de Toronto et la Bibliothèque universitaire Laval, le projet “ notre mémoire en ligne ” se

National Libray of the Czech Republic in Prague, [en ligne]. Adresse URL : <http://www.nkp.cz/digit/digit.htm>.

³⁸ “ L’HTML soutient toute la structure du livre, descendant du niveau supérieur de l’ouvrage vers les niveaux inférieurs, comme le descripteur de la copie numérique l’a établi. Les métadonnées de structures sont balisées. Pour ne pas introduire de nouvelles balises, on a utilisé les balises courantes d’HTML. ” A. KNOLL, *ibid*.

³⁹ “ D’après nous, le langage de base d’Internet constitue l’outil susceptible de prolonger la durée d’accès aux données numériques. En effet, l’accès et la préservation sont, dans ce cadre, étroitement interdépendants et forment un système dans laquelle la solution doit être conçue par harmonisation des deux facteurs ”. A. KNOLL, *ibid*

propose de numériser les collections de Canadiana, retraçant l'histoire du pays, source fortement utilisée par les lecteurs, et reprend pour une part le programme " American memory ". Toutefois, si l'objectif premier reste la mise à disposition plus large d'un corpus bien défini, dont la diffusion sous forme électronique permettra d'économiser les coûts d'entreposage et de manutention à long terme, c'est la question " des coûts, des habitudes d'utilisation et de la viabilité à long terme d'un modèle de collection numérisée "⁴⁰ qui sera l'objet des soins les plus attentifs des concepteurs du projet. En effet, leurs objectifs apparaissent multiples :

- " La conversion d'un ensemble de microfiches au format numérique.
- La transmission électronique via Internet de textes numérisés [...].
- L'examen de l'accès et des questions de récupération reliées au contexte du bilinguisme.
- Le développement du modèle de coût.
- L'analyse de l'utilisation de la collection numérisée et les comparaisons avec l'utilisation de même matériel imprimé et en format microfiche.
- L'essai de la viabilité d'une opération de numérisation coopérative et centralisée ".

Il apparaît, par cet exposé des motifs, que le projet canadien dépasse la simple volonté de mise à disposition des lecteurs d'un fonds retraçant l'histoire du Canada. On peut y voir au contraire une étude grandeur nature des habitudes de lecture de documents numériques dans ce pays, qui décidera pour une grande part de leur avenir et des projets autour de cette technologie dans les prochaines années en Amérique du nord. Dans cette optique, la discussion sur les technologies mises en œuvre a été totalement écartée de la présentation du projet. Au contraire, l'accent a été mis sur les métadonnées qu'il faudra mettre en place pour satisfaire aux habitudes de lecture des usagers. Le projet insiste donc sur le rôle de fournisseur de valeur ajoutée qui sera celui du bibliothécaire des années futures.

⁴⁰ Université Laval. Bibliothèque. (Page consultée le 30 septembre 1998). Notre mémoire en ligne. Description détaillée du projet, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/cihm/ecol/francais/desc-fr.htm>.

D'une certaine manière, le projet de la Bibliothèque nationale de France semble plus classique de par ses aspirations affichées.

c. Un projet à visée encyclopédique : la Bibliothèque nationale de France.

La numérisation à la B.n.F. découle pour une grande part des propositions de P. Cahart et M. Melot de donner à l'utilisation des nouvelles technologies une place plus importante au sein d'une future grande bibliothèque⁴¹. Aujourd'hui, un premier programme de numérisation est achevé, qui a conduit à la mise à disposition du public de 86 000 documents en mode image et quelques documents en mode texte, issus de la base Frantext du Centre national de la Recherche scientifique (C.N.R.S.). Si au départ, un simple accès sur place par les Postes d'Accès à la Bibliothèque Numérique était prévu, l'explosion inattendue de l'Internet a conduit à l'adaptation du produit au web, avec la création de Gallica1 en octobre 1997, site expérimental offrant 5 000 ouvrages en mode image, 250 en mode texte et 7 000 images fixes. Nous en sommes aujourd'hui à Gallica2, site intermédiaire préparant la transition vers Gallica3, qui devrait permettre d'accueillir et de transmettre sur Internet tous les documents numérisés par la B.n.F., libres de droit.

Il est intéressant de noter la volonté d'encyclopédisme⁴² qui a prévalu et conduit au choix des ouvrages⁴³ et le travail en collaboration qui en a découlé. Face aux manques dans les collections de cet établissement, il a fallu faire appel aux ressources de bibliothèques complémentaires, intitulées pôles associés, comme la bibliothèque de l'École polytechnique ou la bibliothèque du Musée de l'Homme. Cet appel à des ressources extérieures constitue un des points forts du programme de numérisation de la B.n.F., qui évite ainsi le cloisonnement disciplinaire qui a longtemps été le sien et

⁴¹ M. RICHARD. "Le programme de numérisation de la bibliothèque de France", dans *Bulletin des bibliothèques de France*, t. 38, n°3, 1993, p. 53.

⁴² Aujourd'hui, on compte 26% d'histoire, 30% de littérature, 15% de sciences, 9% de physique, 8% de sciences politiques, 3% de droit et d'économie et 7% de divers. J.-D. WAGNEUR, "un établissement de référence : la B.n.F. La politique et les réalisations de la B.n.F.", exposé réalisé dans le cadre de la journée d'étude de l'A.C.B. *Les bibliothèques face à la numérisation : enjeux et méthodes. Valenciennes, 6 novembre 1996.*

⁴³ On considère habituellement que 300 000 livres assurent cet encyclopédisme cher à la B.n.F. dans son projet de mettre à disposition en quelque sorte une "bibliothèque de l'honnête homme" du Xxe siècle. J.-D. WAGNEUR, *ibid.*

permet à des bibliothèques de plus petite envergure et moyens de prendre part à cette évolution technologique rapidement.

A ce jour, il est prévu un second programme de numérisation, qui continuera de s'appuyer sur les pôles associés, mais, qui, par souci de lisibilité auprès du public, développera des axes plus thématiques, comme celui de l'utopie, des voyages en Afrique ou des publications des sociétés savantes.

Si le caractère original de l'entreprise de la Bibliothèque nationale de France repose surtout sur la volonté d'encyclopédisme et sur le mode de collaboration qu'il a fallu mettre en œuvre pour tendre à celui-ci, les problèmes techniques se sont posés ponctuellement ou au travers du gigantisme de l'opération.

En définitive, l'étude de ces trois exemples montre la diversité des objectifs que peut offrir la numérisation. Elle montre surtout que les problèmes techniques liés à cette technologie ne sont pas forcément les points les plus importants à résoudre même s'ils apparaissent comme les plus complexes pour une petite bibliothèque ne disposant pas des moyens informatiques des grands établissements présentés ci-dessus. On peut distinguer à nouveau, au travers de ces trois cas, les objectifs de la numérisation : préservation de documents anciens dans le premier cas, ce qui sous-tend la nécessité d'une numérisation de qualité et pérenne, volonté de partager un patrimoine écrit commun, dans le deuxième cas, mais avec le souhait surtout d'étudier les usages de la numérisation pour en tirer des applications notamment dans la création des métadonnées, et enfin volonté d'encyclopédisme dans le dernier, ce qui nécessite la mise en place d'une structure de collaboration.

Les objectifs de Cujas la placeront plutôt dans la première catégorie, mais les enseignements tirés des deux autres exemples, notamment sur les usages, ne sauraient être indifférents lors de la mise en place de cette base numérique.

III. Réflexion autour d'une numérisation de livres anciens à travers l'exemple de Cujas.

La première partie a permis de juger de l'état de la réserve et de ses ouvrages et de comprendre en quoi une numérisation pouvait s'avérer nécessaire. Cette troisième partie intervient après la définition des objectifs de la numérisation, et considèrera, en suivant une chaîne fictive de numérisation, les différents problèmes pouvant survenir à chaque étape et les solutions que l'on peut proposer en fonction des buts de la numérisation et des impératifs liés aux documents anciens.

1. La préparation des documents.

La numérisation ne consiste pas en une simple opération technique pour laquelle la principale difficulté réside en la nécessité de déterminer des critères tels que le mode de traitement des images ou la résolution. Elle réclame au préalable une importante préparation des documents, ne relevant pas simplement de l'ordre du choix de ceux-ci.

On peut distinguer plusieurs étapes de préparation :

• ***Le choix des documents.*** Comme le préconise Mme Bequet, il conviendra de mettre à disposition sous forme numérique les ouvrages fortement consultés dans leur support original⁴⁴. A la détermination d'un corpus strictement intellectuel, panorama de l'histoire du droit occidental pour la bibliothèque Cujas, ou au choix d'un ensemble physiquement déterminé, comme les enluminures de Clairvaux conservées à la B.M. de Troyes, répondra donc obligatoirement une étude poussée sur les usages du fonds proposé à la numérisation, permettant de juger des ouvrages consultés par les lecteurs et des modes d'appropriation de ces derniers. Il conviendra donc de procéder, par exemple, à l'établissement de statistiques de consultation qui permettront de juger du caractère prioritaire de la numérisation de tel ou tel document.

⁴⁴ G. BEQUET. " La numérisation des collections patrimoniales : les enjeux ", intervention du 17 juin 1998 dans le cadre du séminaire organisé par l'E.N.S.S.I.B. *Conservation et mise en valeur des collections patrimoniales.*

Ce travail permettra l'établissement d'une première liste d'ouvrages ou d'images à traiter.

- *L'étude précise des documents choisis.* Elle se divise elle-même en deux grandes parties : l'analyse de l'état physique du document et celle de son accompagnement intellectuel.

Pour l'état physique, il convient de procéder à cette étude pour plusieurs raisons : élimination de documents en trop mauvais état pour faire l'objet d'une numérisation ou éventuelle désignation de ceux qui nécessiteront une restauration préalable ou autre préparation physique. Cette phase permettra en outre de déterminer la volumétrie des pages à traiter, leur nature : texte, gravures ou photographies. Elle servira en outre de base pour la détermination de la valeur du document⁴⁵ et de son état. Ce travail s'avèrera notamment crucial en cas de contestation d'une dégradation par le prestataire.

Plus exigeante encore reste l'analyse du matériel d'accompagnement : il convient en effet que chaque document soit décrit dans une notice bibliographique, dans le cas de livres. Il doit être aussi équipé d'un indice alphanumérique (cote) ou d'un code-barre qui permettra au prestataire de faire le lien entre cette notice, par laquelle se fera la requête, et le document électronique⁴⁶. Cette étape se révèle souvent la plus lourde. Dans le cas de Cujas, elle a par exemple nécessité la reprise systématique de notices souvent erronées. Mais ce travail peut aller encore plus loin : certains fonds, notamment iconographiques, devront être préalablement classés, inventoriés et catalogués. La numérisation de ces fonds peut être évidemment l'occasion de se livrer à cette tâche de façon ponctuelle ; plus souvent, l'importance du travail préalable, s'ajoutant aux activités quotidiennes de la bibliothèque, constitue

⁴⁵ En fonction de celle-ci, le prestataire devra présenter les preuves de la souscription d'une assurance couvrant les pertes ou dégradations commises. G. BEQUET. (Page consultée le 9 novembre 1998). La numérisation des documents patrimoniaux, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/conservation/fr/preventi/guide_dll.htm.

⁴⁶ G. BEQUET, *ibid.*

un frein à la numérisation, comme dans le cas de la bibliothèque des Langues Orientales⁴⁷.

• **Constitution de lots homogènes.** Cette étape se révèle importante car c'est elle qui permettra au prestataire de suivre un rythme continu et soutenu, propre à limiter les dépenses. On pourra distinguer deux types de lots : lot physique, c'est-à-dire entité de livraison, et lot logique, rassemblant des documents aux caractéristiques proches⁴⁸.

• **Préparation physique des documents.** Cette étape consiste par exemple en un nettoyage des documents à numériser⁴⁹ ou éventuellement, dans le cas de livres, d'un démontage de la reliure ou d'un massicotage, comme dans l'exemple de la Bibliothèque nationale de France.

Plus pratiquement, dans le cas de la bibliothèque Cujas, une grille préparatoire, que j'ai réalisée et qui est présentée en annexe 4, devra permettre de faire le point document par document sur les différentes étapes présentées ci-dessus. Elle se comprend à la fois comme un outil de définition et de caractérisation de l'état moyen des ouvrages et de leurs particularités, mais aussi comme un document contractuel envers le prestataire puisqu'il montre à la fois l'état initial du livre et donne les instructions à suivre au moment de la numérisation, en signalant les pages à traiter. Cette grille aura pu être auparavant utilisée par la bibliothèque pour la constitution de lots logiques et homogènes, suivant en cela une chaîne de traitement interne également présentée en annexe 5.

En conclusion, on peut signaler une nouvelle fois l'importance de cette préparation, qui nécessitera souvent la mobilisation des énergies au sein de la bibliothèque et parfois même l'emploi de personnel supplémentaire.

⁴⁷ Exemple rapporté par Mme Buresi, entretien du 29 octobre 1998.

⁴⁸ G. BEQUET, *Ibid.*

⁴⁹ Cela se comprend à la fois pour des raisons de rendu de la numérisation mais aussi d'entretien du matériel. La poussière des fiches de la Bibliothèque interuniversitaire de médecine (B.I.U.M.) encrasse par exemple régulièrement le scanner, source de retard dans la production.

2. La numérisation.

a. Choix du matériel pour numériser.

Différentes possibilités s'offrent pour numériser du texte :

- La scannérisation feuille à feuille, qui présente l'avantage de pouvoir automatiser cette étape. Elle ne s'applique en réalité que pour des documents de bureau.

- La scannérisation à plat ; un programme permet de rectifier la courbure des pages, sans écraser l'ouvrage sur une plaque de verre.

- La scannérisation à livre ouvert, étudiée pour les ouvrages ne pouvant supporter le traitement précédent, actuellement en cours d'élaboration.

- L'utilisation de berceau, qui maintient l'ouvrage à plat, peut être enfin préconisée pour les documents reliés ne pouvant s'ouvrir à plus de 90°.

Comme les documents de Cujas seront conservés après leur numérisation, ce qui exclut le massicotage, le choix se fait essentiellement en fonction de la reliure : il faut en effet que celle-ci permette une ouverture suffisante pour ne pas entraîner de perte d'information. Notons par ailleurs que le coût augmente en fonction du scanner utilisé. Dans le cas des ouvrages de la bibliothèque Cujas, un test réalisé par Jouve a montré que les documents traités à plat n'avaient pu faire l'objet d'une scannérisation sans perte, bien qu'ils aient été sélectionnés en fonction d'une grande possibilité d'ouverture du fait de reliures très endommagées. Il faudra donc envisager la scannérisation à livre ouvert. Une autre solution pourra voir le jour en fonction des ouvrages choisis pour le programme : il faudrait procéder au démontage des reliures de façon à scanner les cahiers à plat. Cette méthode n'est toutefois pas sans risque, notamment de perte⁵⁰ ; elle n'a d'intérêt en outre sur le plan financier que si la restauration s'avère réellement nécessaire. Dans le cas contraire, l'emploi d'un relieur ne représente qu'un surcoût inutile et pas forcément rentable.

⁵⁰ La D.L.L. ne la signale qu'avec toute la prudence nécessaire pour les livres anciens et la déconseille pour les ouvrages uniques. J. DEVILLE. (Page consultée le 9 novembre 1998). La reproduction des documents patrimoniaux. Problématique générale, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/conservation/preventi/guide_dll.htm.

Signalons en outre la possibilité de numériser à partir de microfilms, ce qui est préconisé par la Direction du Livre et de la Lecture. L'absence de ce type de support pour les ouvrages de la réserve de Cujas a fait rejeter cette hypothèse puisqu'il aurait fallu procéder à une campagne de microfilmage préalable, qui aurait endommagé les ouvrages de la même façon⁵¹.

b. Le choix des paramètres de numérisation.

Pour la commodité de l'exposé, nous distinguerons trois grands choix techniques : le mode texte/image, le traitement en bitonal ou en niveau de gris/couleurs et enfin le choix de la résolution à adopter. Il faut bien percevoir que ces différents points s'imbriquent fortement les uns avec les autres et que, dans la réalité le choix d'un critère implique celui des autres.

Mode image ou mode texte ?

Le mode image consiste à numériser un document en mode graphique, c'est-à-dire que le codage informatique retransmet très fidèlement l'image de la page, le style des caractères, la mise en page et les illustrations. Cette méthode ne permet donc en fait qu'une "photographie" de la page, sans possibilité d'intervention ou d'interrogation sur le texte ainsi numérisé et demande en outre une capacité de mémoire jusqu'à huit fois plus élevée. Son intérêt en semble donc limité et il paraît à première vue plus intéressant de numériser le texte en mode caractère.

En effet, ce mode de traitement, transformant l'image en mode point d'un texte imprimé en format texte, le rend compréhensible par la machine et offre alors l'avantage de recherches plein texte et d'intervention sur celui-ci. En outre, en présentant une suite de caractères, il s'avère beaucoup plus facile à compresser. Le seul défaut qu'on peut lui objecter à ce stade de la réflexion reste qu'il ne restitue pas la présentation du texte initial. Dans le cas de la bibliothèque Cujas, l'objectif restant de mettre à disposition de l'utilisateur l'outil le plus performant possible, et au vu

⁵¹ S'est cependant posée la question d'un microfilmage préalable dans un but de pérennisation du support de substitution. Les grandes lignes en seront exposées dans le paragraphe III.6.b.

des avantages du mode texte, cet argument n'aurait pas été décisif et l'on aurait volontiers renoncé à présenter un aspect "livre ancien" qui n'était pas prioritaire dans la définition du projet. Cependant, ce sont les difficultés de passage entre l'image et le texte qui constituent les principaux obstacles au mode texte. La solution la plus simple reste de faire saisir manuellement les textes, optique retenue par O. Bogros à la B.M. de Lisieux. On conçoit aisément que ce ne peut être valable que pour un nombre restreint de pages, pour un usage ponctuel, et ce pour des raisons évidentes de coût en main d'œuvre.

La deuxième solution, aujourd'hui largement répandue, consiste en l'utilisation d'un logiciel de reconnaissance optique des caractères qui va agir en 5 étapes :

- identification des blocs texte par rapport aux blocs illustrations,
- reconnaissance des caractères,
- identification ou reconnaissance du mot,
- correction,
- mise en forme de la sortie.

Toutefois, malgré les améliorations constantes des logiciels, de nombreux facteurs en affectent l'exactitude : résolution du scanner, génération et dimension de l'original, reliure, ...⁵² qui font qu'en temps normal, le degré de fiabilité d'un logiciel de reconnaissance est de l'ordre de plus de 99%, soit 1 à 10 caractères erronés tous les 1 000, taux importants quand l'on travaille sur les 100 000 pages du projet de Cujas...

A ces problèmes d'ordre général s'ajoutent ceux plus particuliers des textes anciens : une typographie ancienne, des lettres couplées, des caractères gothiques, qui entraîneront forcément un taux plus élevé d'erreurs et la nécessité d'un contrôle plus poussé et donc plus coûteux. On doit également y ajouter le papier jauni par les ans, parsemé de tâches de roussissures ou d'humidité : le test réalisé par Jouve sur les ouvrages de Cujas montrait, en bitonal, que le contraste noir/blanc peu marqué parfois entre le texte et le papier entraînerait un accroissement des erreurs du logiciel de

⁵² S. HAIGH. (Page consultée le 9 juillet 1998). La reconnaissance optique de caractères (R.O.C.) en tant que technologie de numérisation, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/pubs/netnotes/fnotes37.htm>

Reconnaissance optique de caractères (R.O.C.). L'empâtage des caractères, tel qu'il apparaît en bitonal, constituerait une difficulté de plus pour la reconnaissance de ceux-ci. Les conseils donnés au Conservatoire national des Arts et Métiers (C.N.A.M.) nous ont donc dissuadé d'employer cette méthode, d'autant que son coût élevé, 20 F⁵³ par page, risquait de s'ajouter à un très important taux d'erreurs.

Dans le cas de livres anciens, le choix du mode image s'avère donc plus que jamais incontournable. Un tableau présenté en annexe 8 permettra de résumer, de façon synthétique et visuelle, les avantages et inconvénients des deux modes de traitement que nous venons de présenter.

Choix du mode de traitement.

Nous allons distinguer ici trois grands modes de traitement, bitonal, par niveaux de gris et par niveaux de couleurs, et nous appuyer dans notre réflexion sur un test portant sur onze pages significatives de la réserve, réalisé par la société Jouve, pour tâcher de présenter la solution la plus pertinente dans le cas des ouvrages de Cujas. Une sélection de ces pages numérisées est présentée en annexe. Nous traiterons par la même occasion du format du fichier et du mode de compression, dont le choix découle souvent directement du mode de traitement retenu.

Nous avons, dans tout ce paragraphe, écarté volontairement les formats propriétaires, appelés ainsi car leur utilisation nécessite le versement de droits à la société qui peut, par ailleurs, mettre en libre accès les outils de lecture : P.D.F. (Adobe) ou PCD (Kodak). Leur fréquence d'utilisation en ont fait des normes de fait, mais ils restent toutefois soumis au bon vouloir du fabricant et peuvent, en conséquence, présenter des risques pour la conservation des données numériques, point essentiel quand il s'agit de documents anciens.

⁵³ Ce tarif, ainsi que les suivants, ont été donnés par la société Jouve SI. suite à un test de numérisation réalisé dans leur entreprise. Ils s'entendent pour le faible nombre de pages qui leur avaient été confiés et ne reflètent pas la réalité des coûts à envisager pour une opération plus complexe. Ces tarifs sont donc surtout donnés ici pour fournir un ordre d'idée.

• Le bitonal. Cette technique consiste en un codage de chaque pixel de l'image sur un bit, c'est-à-dire que seuls le noir et le blanc apparaîtront à l'écran. Cette solution s'avère extrêmement économique, aussi bien sur le plan de la mémoire, d'où un transfert rapide, que financièrement (autour de 4 F la page). Elle offre en outre un bon contraste noir/blanc, qui la désigne souvent comme le mode image préparatoire au passage d'un logiciel de R.O.C. Ces qualités en font un mode de traitement particulièrement recommandé pour le texte, y compris pour l'archivage, choix qui a été fait d'ailleurs par la Bibliothèque du Congrès⁵⁴.

Cette technique comporte, de par ce codage blanc/noir, un risque non négligeable pour une numérisation de qualité d'ouvrages anciens. Les tirages, fournis en annexe des pages numérisées par Jouve, permettent d'illustrer ce propos. Il était en effet à craindre que les taches de roussissures, les papiers jaunis ou brunis, les taches d'humidité qui parsèment les ouvrages anciens ne soient interprétés par le scanner comme des points à traduire en noir, rendant ainsi illisible le texte écrit sur cette page. Le test a permis de nuancer ce propos : les tâches d'humidité prononcées apparaissent effectivement, sous la forme d'une longue traînée noire, alors que celles de faible densité sont interprétées comme du blanc et disparaissent donc. Le papier brun ou jauni ne s'est pas avéré incompatible avec le bitonal, dans la mesure où l'encre restait suffisamment noire pour permettre le contraste, mais entraîne un problème de lissage des caractères plus important, tout en restant parfaitement lisible.

En définitive, le bitonal peut être envisagé pour les livres de la réserve, même s'il fait perdre le côté "document ancien" par la perte de l'aspect du papier. Cette contrainte ne s'avère pas gênante pour la bibliothèque Cujas, qui veut mettre un outil à disposition et non faire une exposition des documents de la réserve. Le bitonal exige une grande attention et un examen approfondi des ouvrages. Le test a permis de rendre compte des particularités de papier ou d'encre entraînant un mauvais rendu : il conviendra, sur la liste des ouvrages sélectionnés, de voir ce qui relève de l'état moyen et ce qui relève d'exceptions, en comparant avec les pages-repères, pour voir dans

⁵⁴ C. FLEISCHHAUER. (Page consultée le 30 septembre 1998). Digital formats for content

quelle mesure le choix du bitonal ne pénalisera pas trop la numérisation de la bibliothèque. Pour les pages pour lesquelles ce choix posera un problème, trois solutions se présenteront : choix du bitonal tramé⁵⁵, rejet de l'ouvrage entier ou passage outre, les deux dernières solutions dépendant directement du nombre de pages altérées.

Le bitonal entraînera, dans la pratique, le choix d'un format T.I.F.F. et d'une compression U.I.T.-IV⁵⁶, du type utilisé pour la télécopie, et particulièrement efficace du fait des caractéristiques du bitonal. On préférera une compression J.B.I.G. dans le cas du bitonal tramé.

• Niveau de gris. Ce mode de traitement consiste à traiter l'image grâce à une palette de gris, dont le nombre dépendra du nombre de bits employés par pixel : un codage sur 8 bits donnera ainsi une palette de 256 gris. Cette technique, et le test l'a confirmé sans grande surprise, a permis un bon rendu de toutes les pages proposées, indépendamment des particularités de chacune. Son rendu esthétique, maintenant l'aspect livre ancien, et cette sécurité dans le rendu des ouvrages anciens, l'ont fait adopter dans la quasi-majorité des numérisations de documents patrimoniaux, que ce soit par la Bibliothèque nationale de France, le Conservatoire national des Arts et Métiers ou la Bibliothèque du Congrès. Il présente cependant quelques inconvénients : un coût plus élevé que le bitonal (autour de 6 F/page), et surtout un plus important volume en mémoire, environ huit fois plus. Il conviendra donc, si l'intention de la bibliothèque est de la diffuser par les réseaux, d'en proposer une version compressée. On traitera des fichiers au format T.I.F.F. ou G.I.F., en général, grâce à la compression J.P.E.G., norme reconnue et donc sûre. Il conviendra de doser le taux de compression pour allier efficacité de transmission et respect de la qualité de l'image. Par exemple, la Bibliothèque nationale de France a adopté la compression

reproductions, [en ligne]. Adresse URL : <http://memory.loc.gov/ammem/formats.html>.

⁵⁵ La mise en place d'un tramage, fin réseau de lignes obliques, donne l'impression " niveau de gris " pour des documents numérisés en bitonal. Toutefois, ce tramage s'avère gênant à l'œil et ne donne pas forcément de meilleur résultat.

⁵⁶ Anciennement CCIT-groupe IV.

JPEG 1:10. Notons toutefois que cette compression importante se fait avec perte d'information et donc de qualité de l'image. Il sera donc recommandé à une bibliothèque y ayant recours et tenant à une certaine pérennité de ses enregistrements numériques d'exiger du prestataire une version non compressée pour la conservation et une version compressée pour la diffusion⁵⁷.

- Niveau de couleur. Cette méthode repose sur le même principe que le niveau de gris, mis à part qu'elle superpose en quelque sorte trois numérisations en trois couleurs différentes : rouge, vert, bleu. Ce mode de traitement, même à une plus faible résolution que les 300 dpi sur lesquels s'est fondé le test, a sans conteste donné les meilleurs résultats, alliant contraste et effet esthétique, lisibilité et rendu des détails. Instinctivement, cette méthode rallierait sans doute les suffrages. Elle présente malgré tout deux défauts importants : son coût (autour de 9F la page) et son volume en mémoire, même après une compression qui s'avère souvent plus efficace que pour un fichier identique en niveau de gris. Avant de choisir ce type de numérisation, il faut donc se demander réellement les enjeux du projet et s'interroger sur la nature des documents. Pour des documents sans illustration, ou juste quelques gravures, comme dans le cas de Cujas, cette solution apparaît difficilement recevable alors que la B.M. de Troyes l'a choisie pour son projet, plus iconographique car il s'agit des livres de la bibliothèque bleue⁵⁸. Enfin, si Cujas peut envisager de se pencher sur la numérisation en couleur, pour des raisons pragmatiques de coût, ce ne pourra sans doute être que pour quelques pages plus publicitaires, pour illustrer la richesse de la réserve sur le site Web par exemple, mais non dans le cas d'un programme de 100 000 pages !

⁵⁷ G. BEQUET, *ibid.*

⁵⁸ Littérature de colportage, imprimée sur du papier de qualité médiocre et protégé d'une couverture bleue, d'où son nom. Fabriquée essentiellement à Troyes, cette collection s'adressait à un public majoritairement illettré et était donc de ce fait abondamment illustrée.

Signalons tout de même que, pour cette numérisation couleur, on associera le plus souvent un format T.I.F.F. ou G.I.F., également associés à une compression J.P.E.G⁵⁹.

Le tableau présenté ci-dessous permet de faire le point sur les avantages et les contraintes de chaque méthode.

<i>Format</i>	<i>Bitonal</i>	<i>Niveau de gris</i>	<i>Niveau de couleur</i>
<i>Avantages</i>	<ul style="list-style-type: none"> • Coût • Contraste noir/blanc • Faible volume en mémoire • Rapidité d'accès • Qualité de lecture et d'image suffisante quand la page est saine • Disparition des taches d'humidité peu prononcées • Bonne qualité pour la gravure sur bois • Facilité d'impression • Compression 	<ul style="list-style-type: none"> • Lisibilité • Rendu de l'aspect initial • Finesse dans le rendu des gravures sur cuivre et bois • Pas de problèmes de lecture en cas d'humidité, de papier bruni ou jauni 	<ul style="list-style-type: none"> • Lisibilité • Rendu de l'aspect initial • Rendu des détails • Rendu de bonne qualité quelles que soient les caractéristiques de l'original
<i>Inconvénients</i>	<ul style="list-style-type: none"> • Problème de lissage des caractères quand le papier est jauni ou bruni, d'où une moins bonne lecture • Mauvais rendu esthétique • Perte du caractère ancien de l'ouvrage • Incompatibilité avec de fortes taches d'humidité • Moindre rendu pour la gravure sur cuivre 	<ul style="list-style-type: none"> • Coût • Compression, d'où problème de conservation • Fort volume en mémoire • Papier au fond grisé sur encre peu foncée posant des problèmes de contraste 	<ul style="list-style-type: none"> • Coût • Lenteur d'accès • Nécessité de compression pour la diffusion, d'où un problème pour la conservation • Fort volume en mémoire
<i>Exemples d'utilisation</i>	<ul style="list-style-type: none"> • Bibliothèque du Congrès pour les ouvrages fin XIX-XX^e siècles 	<ul style="list-style-type: none"> • B.n.F. • B.n. du Québec • Bibliothèque du Congrès pour les ouvrages anciens • C.N.A.M. • Université du Michigan 	<ul style="list-style-type: none"> • B.M. de Troyes

⁵⁹ G. BEQUET, B. CARBONE, P. DEBRIEN et al. (Page consultée le 5 mai 1998). La bibliothèque multimédia, [en ligne]. Adresse URL : <http://www.culture.gouv.fr/culture/mrt/numerisation/fr/dll.htm>.

Choix de la résolution.

De la résolution, nombre de pixels par unité de longueur, dépendra la qualité de l'image numérisée, l'importance du détail, mais aussi le volume du fichier. L'enjeu apparaît donc de concilier un niveau de résolution suffisamment important pour assurer la conservation des documents électroniques, et suffisamment réduit pour en permettre la diffusion. Dans le cas d'une numérisation de textes, il semblerait qu'une résolution de 300 dpi soit un bon compromis, choisi notamment par la Bibliothèque du Congrès⁶⁰. Une autre solution consiste à effectuer une numérisation dans une très haute résolution (jusqu'à 1200 dpi, mais le plus souvent 600 dpi) et à en tirer des produits destinés à la diffusion à plus faible résolution. Les ouvrages de Cujas, anciens certes, mais d'une valeur bibliophilique réduite si on les prend isolément, ne le nécessiteront sans doute pas.

c. Le contrôle qualité.

Cette étape, souvent négligée, s'avère primordiale de l'avis des personnes responsables d'un projet de numérisation aujourd'hui parvenu à son terme. Pour M. Coronio, responsable de la bibliothèque de l'École des Ponts et Chaussées, par exemple, le renforcement du contrôle sera sans doute le point sur lequel il fera porter l'essentiel de la réflexion de son équipe pour le prochain programme.

Les enjeux du contrôle de l'opération de numérisation se révèlent en effet extrêmement importants, dans la mesure où il décide en partie de la réussite du projet et peuvent éviter, à l'avenir, une seconde manipulation des ouvrages, souvent néfaste, en particulier dans le cas de documents anciens comme ceux de Cujas. Le prestataire chargé de l'opération de numérisation proprement dite mettra en place au sein de son entreprise des moyens de contrôle de la production d'autant plus forts qu'il se sentira lié par les vérifications que pourra exercer l'établissement commanditaire. Il convient donc d'avoir des idées précises des moyens en hommes et en matériel que l'on pourra mettre en œuvre, de le préciser dans le cahier des charges

⁶⁰ C. FLEISCHHAUER, *ibid.*

pour que le prestataire puisse en prévoir les répercussions sur l'organisation de son travail et de mentionner la structure des équipes responsables du contrôle et le matériel à disposition.

Plus les précautions que devra prendre le prestataire pour éviter le rejet des lots par la bibliothèque seront lourdes, plus cela se répercutera sur les tarifs qu'il demandera. Il apparaît, de toute manière, que la bibliothèque doit considérer la mise en place d'une base de documents numérisés comme celle d'un système intégré de gestion de bibliothèque, et ne payer les prestations de la société de sous-traitance qu'après vérification du service rendu, soit après un contrôle extrêmement précis qui semble normal dans le second cas alors qu'il est souvent négligé dans le premier.

En conséquence, l'exposé des modalités de contrôle du travail de numérisation constitue un paragraphe extrêmement important du cahier des charges techniques particulières. Il convient de se montrer très précis pour éviter par la suite toute contestation du prestataire. Les lenteurs de règlement du litige retarderaient d'autant la mise en place des documents numériques et le retour en magasins de livres à la fois fragiles et beaucoup consultés, dans le cas de Cujas.

Dans le cas de la Bibliothèque nationale de France, le contrôle s'est fait en deux temps :

- Contrôle automatique, réalisé systématiquement par l'ordinateur, afin de vérifier les formats informatiques, la conformité des fichiers par rapport à la mire de contrôle définie au début de l'opération et la présence effective de ces fichiers;

- Contrôle humain, réalisé par des vacataires à mi-temps, sur un nombre restreint de bandes, désignées quantitativement et qualitativement en fonction de la norme relative aux échantillonnages. Deux types d'erreurs sont alors distinguées :

⇒ les erreurs rédhibitoires entraînant le rejet pur et simple du lot : ouvrages incomplets (avec une attention particulière sur la présence des pages importantes, comme page de titre), les informations tronquées, les paginations incohérentes, les références bibliographiques incohérentes, la visualisation impossible.

⇒ Les erreurs cumulatives : le rejet du lot se fait à partir d'un certain nombre, décidé préalablement : impressions de travers (il convient alors de fixer un angle maximum toléré), visibilité insuffisante, données de références bibliographiques erronées⁶¹.

Dans le fond, ce type de contrôle, tel que défini par la Bibliothèque nationale de France, reste valable pour tout type de numérisation ; dans la pratique, il est à craindre que la bibliothèque Cujas, comme d'autres, ne disposera pas de moyens humains et matériels suffisant pour mettre en place un contrôle aussi poussé au sein de son établissement.

Une solution plus réaliste sera alors de suivre l'exemple proposé par T. Delcourt à la B.M. de Troyes, consistant à préciser dans le cahier des charges que la vérification se déroulera dans les locaux du prestataire. Selon une périodicité fixée entre le titulaire et la bibliothèque, elle se fera en présence du chef de projet utilisateur et du conducteur des travaux⁶².

Cette méthode, proche de celle adoptée par les Ponts et Chaussées, semble constituer un bon compromis entre les moyens réduits d'une bibliothèque et les nécessités de contrôle de la qualité de la numérisation, gages d'une réussite du projet.

3. Les outils d'accès.

On peut distinguer deux types d'outils d'accès : ceux permettant la navigation au sein du document, c'est-à-dire la table de correspondance, et ceux répondant à une démarche intellectuelle de recherche d'un titre ou d'une page spécifique, à savoir les notices et les tables des matières, outils conçus voire réalisés par la bibliothèque.

a. La création de la table de correspondance.

La constitution de la table de correspondance s'avère un moment incontournable dans le processus de numérisation d'un ouvrage. Réalisée par le

⁶¹ T. BOUCHER. *La numérisation des documents imprimés à la Bibliothèque nationale de France*. Mémoire d'étude, sous la direction de T. Lafouge, E.N.S.S.I.B., DCB 1994, n°28, p.28-29.

prestataire, dans le cas où la bibliothèque a fait appel à une société de sous-traitance, elle permettra en effet la circulation au sein du document numérisé de façon transparente pour l'utilisateur. Son rôle consiste à faire le lien entre la page du document imprimé original et le fichier électronique correspondant. Il faut donc, pour chaque ouvrage, créer une table, fichier particulier établissant le lien entre un numéro de page de l'ouvrage physique et son numéro d'ordre de numérisation. Cette table se fait en général automatiquement. Il convient de signaler qu'elle ne constitue pas une étape inutile dans la mesure où les deux numéros ne coïncident pas toujours. Par exemple, la page de titre peut ne pas être paginée dans la réalité, alors qu'elle fera bien l'objet d'un fichier numérique. S'institue ainsi un décalage réalité-numérique que la table de correspondance vient corriger.

A ces problèmes d'ordre général sur la constitution de cette table, s'ajoutent certaines difficultés spécifiques aux ouvrages anciens, dont la réserve de Cujas représente un exemple. L'étude des documents prévus pour la numérisation a révélé des difficultés de plusieurs types ; on a constaté ainsi :

● **L'existence d'une pagination par folio, page ou colonne.** La pagination par colonne a pu d'ailleurs être compliquée par une présentation, relevant d'une impression encore fortement influencée par la mise en page manuscrite, consistant à entourer un texte, constituant un bloc central, par de la glose, disposée elle en deux colonnes. Dans ce dernier cas, la pagination se fait souvent par colonne.

● **Une pagination arabe et romaine.** On assiste souvent à une succession de celles-ci, surtout dans le cas de recueils. A la difficulté qu'occasionne la présence de ce double type de pagination, s'ajoute celui de leur répétition. Pour un même volume représentant parfois en même temps une unité physique et intellectuelle, on pourra donc avoir plusieurs pages " 2 ", par exemple, mais aussi plusieurs pages " II ".

● **L'existence ou non de pagination.** Dans le cas de livres très anciens, comme ceux du XVI^e siècle, les signatures n'ont pas encore laissée place à la pagination ou foliotation moderne. Pour des bibliophiles, ces indications permettent

⁶² C.C.T.P. de la B.M. de Troyes, p.9.

de se repérer au sein du livre mais restent sans doute transparentes pour le prestataire, pas toujours compétent en livre ancien. Sans parler de cet exemple, qui reste minoritaire parmi les ouvrages de Cujas mais qu'il fallait signaler, certains livres comportent à la fois des pages numérotées et d'autres qui le sont pas, en début et fin du livre, comme celles présentant le privilège, mais parfois aussi au cœur de l'ouvrage : il convient donc de les signaler au prestataire chargé de la numérisation, qui n'y prêterait peut-être pas garde.

Ces quelques exemples que nous soulignons ici à partir de l'analyse des ouvrages de Cujas permettent de montrer que, dans le cas de livres anciens, chaque document présente ses propres particularités en matière de pagination. Il s'avère donc impossible de "plaquer" sur un ensemble de monographies un modèle présumé de table de correspondance.

Face à ces difficultés, le Conservatoire national des Arts et Métiers (C.N.A.M.) et la Bibliothèque nationale de France ont apporté deux réponses différentes.

Le C.N.A.M. a traité des ouvrages du XVIII^e siècle de petit format ; il a été demandé au prestataire de réaliser des vues de deux pages. Chaque vue est enregistrée dans un fichier doté d'un numéro d'identification propre. Sera ensuite créé un programme qui découpera chaque fichier en deux pour reconstituer les pages, les "nettoiera", changera le format et permettra à l'opérateur de faire la correspondance avec la table des matières, entrée en mode texte. Cette méthode présente l'inconvénient de ne pas prendre en compte les pages non présentes dans la table des matières, qui ne seront donc accessibles que par la fonction "page suivante"⁶³.

Pour la B.n.F., où chaque page était contenue dans un fichier distinct, la table de correspondance comportait trois zones distinctes : le numéro du fichier électronique, le numéro de page toujours exprimé en chiffres arabes, le type de pagination (arabe, romain, non paginé ...)⁶⁴.

⁶³ Réponse de P. Cubaud, maître de conférence en informatique au C.N.A.M., par e-mail.

⁶⁴ Réponse de G. Cathaly, service numérisation de la B.n.F.

En définitive, la création de la table de correspondance, surtout pour le livre ancien, s'avère souvent une opération délicate. Même si elle reste le travail du prestataire dans la plupart des cas, il convient toutefois de signaler à celui-ci au préalable les particularités de chaque ouvrage, pour qu'il puisse adapter son modèle et réaliser ainsi une table de correspondance efficace.

b. L'accès par la notice.

Le mode d'accès par la notice doit être considéré comme le premier, à la fois dans la chronologie de l'utilisation de la base numérisée, mais aussi intellectuellement pour un lecteur formé à la consultation d'un catalogue. Il importera donc, comme nous l'avons vu, d'apporter tous ses soins à la vérification du fonctionnement du lien entre la notice et les documents électroniques et à l'exactitude de la saisie de celle-ci.

Pour l'exemple de Cujas, un intérêt particulier a été consacré à ces notices. Nous avons vu que celles-ci se résumaient à une brève description sur l'une des fiches de format carte à jouer du catalogue des ouvrages entrés avant 1952. Un premier travail a donc consisté, suivant l'exemple de la bibliothèque du Conservatoire national des Arts et Métiers, à les reprendre entièrement, les corriger et à les enrichir éventuellement. Il faut noter d'ailleurs que cet aspect ne doit pas être perçu comme caractéristique de la situation de Cujas, et prendre en compte que de nombreuses bibliothèques n'ont pas eu le temps ou les moyens de rétroconvertir leurs vieux fichiers, voire de cataloguer leurs fonds anciens. Une étude poussée des notices reste donc inévitable et particulièrement cruciale dans le cas de collections patrimoniales.

A Cujas, cette tâche s'est effectuée manuellement, pour ne pas compliquer la future rétroconversion en créant des doublons, même s'il aurait semblé plus logique d'en profiter pour les entrer dans la base informatisée de la bibliothèque. Cette pratique nécessitera donc de faire saisir les notices catalographiques par le prestataire : pour des raisons de coût, il a donc été convenu de réaliser des notices ne comprenant pas tous les champs habituels pour le catalogage du livre ancien mais limitées au minimum, c'est-à-dire les champs que l'on considère comme indexables : auteur, titre,

mentions d'édition, indexation matière et cote. Une attention particulière a été apportée en outre aux auteurs secondaires et aux titres.

Dans la plupart des cas en effet, les auteurs secondaires n'étaient pas mentionnés dans les fiches. Or, dans le cas des ouvrages de Cujas, cet aspect a semblé très intéressant pour obtenir une réponse plus fine à l'interrogation du lecteur. On peut prendre l'exemple du *Corpus juris civilis*. L'auteur en est Justinien, et il convient donc de le mentionner en auteur principal, mais il ne faut pas oublier les traducteurs, commentateurs, ... Les usages des historiens du droit prouvent qu'ils consultent les différentes éditions, surtout en fonction de l'intérêt du glossateur : il convient donc de soigner ses accès.

Le même soin a été apporté, grâce à un catalogage à niveau, aux titres des ouvrages contenus dans un recueil pour favoriser l'accès.

Dans la pratique, le prestataire aura donc à charge de saisir ces notices bibliographiques enrichies et corrigées, vraisemblablement en format Marc, et de constituer le programme qui permettra l'indexation des différents champs interrogeables et celui qui fera le lien entre les notices et le document électronique.

Dans le cas de Cujas, l'absence d'informatisation du catalogue d'avant 1952 a empêché d'envisager une consultation des documents électroniques à partir du catalogue. La rétroconversion est programmée sur le même plan quadriennal que la numérisation, sans que l'on puisse dire lequel des deux projets sera d'abord réalisé. Toutefois, de nombreuses bibliothèques, à l'exemple de la B.M. de Valenciennes ou de la bibliothèque nationale du Québec, ont fait le choix de lier l'interrogation des documents numériques au catalogue informatisé, option qui pourra être envisagée à l'avenir à Cujas. Dans ce cas, le lien se fait entre la notice catalographique d'un document numérisé et ce dernier, par le remplissage du champ 856 du format Marc. Cette méthode nécessite cependant une forte capacité en mémoire quand on sait que la majorité des ouvrages numérisés comportent un minimum de 15 Mo⁶⁵.

⁶⁵ Réponse au questionnaire sur biblio-fr de M. Y-A. Lacroix, bibliothèque nationale du Québec.

c. L'accès par la table des matières.

L'accès par la notice constitue le premier moyen de retrouver facilement le texte que l'on veut consulter. Toutefois, dans le cas de Cujas, il nous a semblé un peu réducteur et mal commode, pour des corpus de parfois 10 à 15 volumes. Certes, un catalogage à niveau permettait déjà de repérer des œuvres précises, au sein d'un recueil par exemple, mais il ne résolvait pas le principal problème, celui d'une navigation optimisée au sein d'un texte de plusieurs milliers de pages. De plus, se pose également le problème de l'accès à distance au texte : il convient d'éviter à l'internaute le rapatriement long et coûteux de nombreuses pages inutiles et de lui permettre de mieux cibler sa demande.

Pour pallier cet inconvénient, il faut souligner le choix qui a été fait aussi bien par la Bibliothèque nationale de France que par le Conservatoire national des Arts et Métiers⁶⁶ de numériser en mode texte les tables des matières. Pour les mêmes raisons qu'expliquées plus haut, le passage d'un logiciel de reconnaissance optique de caractères n'a pu être envisagé et il a donc fallu se résoudre dans les deux cas à la saisie manuelle des tables des matières.

Ce travail a été confié à une entreprise de sous-traitance dans le cas de la B.n.F., avec l'option de traiter les index en cas d'absence ou de déficience de la table. Le C.N.A.M., disposant de moyens plus réduits pour un projet plus modeste, a préféré confier cette tâche à un vacataire, ce qui présentait l'avantage indéniable de contrôler et corriger le texte saisi en H.T.M.L. au fur et à mesure. Dans certains cas, il a été fait le choix de créer de toutes pièces les tables manquantes.

Pour les deux bibliothèques, la saisie des tables présentait un double avantage :

- Rendre possible l'interrogation par mots-clés sur le contenu de la table des matières, en indexant celles-ci. Ce système permet ainsi une interrogation plus large que par les mots du titre.
- Rendre possible surtout un accès direct à la page intéressante, sans passer par la table de correspondance, en mettant en place des liens directs entre la table des

⁶⁶ Entretien du 8 octobre 1998 avec le conservateur chargé du fonds ancien.

matières et les pages concernées (unité intellectuelle, début de chapitre, début de subdivision selon la finesse choisie).

Il apparaît évident que ce système offrira une forte valeur ajoutée aux documents numérisés en allégeant le travail fastidieux de feuilletage au lecteur. Il est d'ailleurs intéressant de constater qu'une enquête sur les usagers des textes juridiques de Gallica laissait apparaître que les lecteurs se contentaient de consulter la table des matières sans forcément s'intéresser au texte proprement dit⁶⁷.

Signalons toutefois que cette méthode, toute satisfaisante intellectuellement qu'elle soit, n'est pas sans poser des difficultés. La saisie manuelle de pages entières a un coût important qu'il convient de ne pas négliger. Une bibliothèque comme Cujas, qui souhaite mettre à disposition un outil performant, préférera sans doute sacrifier le nombre des ouvrages à la saisie des tables de ceux qui auront été définitivement retenus. Il faut ensuite prévoir des structures de contrôle des tables saisies, dont il aura fallu auparavant préciser pour chacune le degré de précision, afin d'en vérifier la complétude et la construction, dans le cas de tables formées *ex nihilo*.

A Cujas, pour les ouvrages sélectionnés prioritairement pour la numérisation, l'effort de saisie porterait sur 670 pages et celui de création sur environ cinq titres. Toutefois, en ce qui concerne les tables existantes, on note une grande variété dans la présentation et le degré de précision. Il sera sans doute en conséquence possible d'en faire diminuer l'ampleur sans pénaliser la consultation.

Il me semble en définitive que les avantages de la création d'une table des matières l'emportent largement sur les inconvénients, quand il s'agit de documents de travail que l'on souhaite mettre en valeur. De plus, il convient de souligner que cet aspect de la numérisation et de la réflexion autour des accès créés illustre parfaitement le nouveau rôle que devra assumer le bibliothécaire confronté aux nouvelles technologies de l'information : celui de créateur de valeur ajoutée autour des documents qu'il met à disposition.

⁶⁷ Entretien téléphonique avec M. Dussert, responsable des acquisitions juridiques pour la numérisation à la B.n.F.

4. Le stockage.

Nous avons ici choisi de présenter plus particulièrement les supports de stockage et ceux de diffusion, à l'exclusion de ceux d'archivage. Dans le cas de Cujas, comme nous le verrons, on envisagera sans doute le même support pour le stockage et pour l'archivage, même si leur contenu pourra différer pour des raisons de pérennité de l'information numérique. Nous étudierons ici plus précisément la distinction qu'il convient d'adopter entre support de diffusion et support de stockage pour allier le confort de communication des informations et une bonne conservation des données. Un tableau présenté en annexe 9 synthétisera les développements suivants.

a. Support de stockage.

Pour stocker les données numériques et leur assurer une conservation de bonne qualité, plusieurs solutions se présentent :

- **Bandes magnétiques**, à l'exemple de cassettes DAT. Réinscriptibles et recommandées pour la sauvegarde régulières de données, elles peuvent éventuellement servir de support de livraison mais n'offrent pas de garanties suffisantes pour la conservation⁶⁸.

- **Disques Optiques Numériques (DON-WORM)**. Leur forte capacité de stockage (10 à 25 Go) en font des supports recommandés pour la conservation des données sur plusieurs années pour les textes, sons ou images et ont été choisis par exemple par la B.n.F. Ils nécessitent toutefois un matériel de lecture et d'écriture spécifique.

- **CD-WORM (ou CD-R)**. D'une capacité de 650 Mo, soit l'équivalent de 15 000 images, ce support offre l'avantage de coûter peu cher, d'être très répandu et de fonctionner avec un lecteur de cédérom, matériel courant en bibliothèque.

⁶⁸ Pour information, la B.n.F. vient de passer commande auprès de la société Digipress de 3 440 C.D. en verre. Les 86 000 pages numérisées, représentant 2 téraoctets de mémoire seront ainsi transférés de bandes magnétiques DAT vers ce nouveau support, le century disque, réputé plus résistant que le simple C.D. en plastique. Message de J. Laclotte diffusé sur biblio-fr le 11 janvier 1999.

• **DVD-R.** De même dimension que le cédérom, il offre de plus importantes capacités de stockage : environ 7,9 Go. A l'avenir, il deviendra certainement le support le plus couramment utilisé. Aujourd'hui toutefois, les discussions menées entre fabricants n'ont pas permis la définition d'une norme pour ce produit. Il convient donc de rester prudent quant à son utilisation.

En définitive, le choix de Cujas pour son support de stockage se portera sans doute sur le cédérom, choisi d'ailleurs par l'essentiel des bibliothèques ayant mené un projet similaire en France, même si cela signifie un nombre important de disques. On pourra par contre envisager, dans les années à venir, de transférer les données vers les DVD-Rom quand ceux-ci présenteront des garanties suffisantes.

Support solide, le cédérom doit tout de même faire l'objet de quelques précautions : conservation dans un milieu d'une température entre 5 et 20°, taux d'humidité relative entre 20 et 50%, pression atmosphérique entre 75 et 105 kPa. Il faudra veiller à les conserver à l'abri de leur jaquette, à faire attention aux produits de nettoyage et ne pas les exposer à la lumière⁶⁹.

b. Support de diffusion.

Le choix du cédérom comme stockage implique un changement de support pour l'exploitation des données numériques. Il faut en effet envisager un nombre très important de ces disques optiques pour renfermer les 100 000 pages qui seront numérisées par Cujas. Leur consultation exigera donc un important dispositif de tours de cédéroms ou de juke-box, qui peuvent s'avérer très lourd à gérer. De plus, l'utilisation intensive de ces disques s'avère toujours risquée même s'ils ne sont pas fragiles. On est en effet toujours à la merci d'une rayure ou de la perte de la couche de vernis protectrice entraînant un risque de détérioration des enregistrements. Enfin, l'utilisation des cédéroms pour une consultation en ligne, peut augmenter le délai d'obtention des informations car "cette technique manque parfois de souplesse,

⁶⁹ G. BEQUET, *ibid.*

même si elle a énormément progressé. L'enregistrement n'est pas toujours intuitif, et les vitesses de transfert sont plutôt faibles ”⁷⁰.

Ces raisons expliquent que dans tous les exemples de bibliothèques contactées, il ait été fait clairement une distinction entre le support de stockage et le support d'exploitation : les données conservées sur cédérom sont ainsi immédiatement transférées sur le disque dur d'un serveur, dont la vitesse et la capacité en mémoire, ont été étudiées pour gérer cette masse importante d'information. Dans le cas de la Bibliothèque nationale du Québec, les 40 giga-octets que représentent les 360 000 pages numérisées sont ainsi accueillies sur le serveur Web, un IBM Risc 6000⁷¹.

c. Catalogage des données numériques.

L'importance prise ces dernières années par les documents électroniques, en particulier depuis l'expansion de l'Internet, pose au bibliothécaire la question de l'archivage mais aussi du catalogage de ces données. Les documents numérisés doivent-ils, dans ce contexte, faire l'objet d'une notice spécifique ? La question se pose dans le cas de Cujas, comme dans d'autres.

L'examen du projet de numérisation de microfilms mené à l'université de Yale répond par la négative à cette question, considérant que dans ce cas, le document numérisé n'est qu'une reproduction de l'original, même si elle n'est pas “ digne de foi ”, puisque comportant des informations supplémentaires (métadonnées) ou ne reproduisant pas exactement l'original : “ the digital image file is an image-based representation of an original document and as such we treat it as a reproduction ”⁷².

⁷⁰ F. FROSSARD. “ Stockage amovible : à chaque technologie son mode d'utilisation ”, dans *Informatique*, n°1481, 16 janvier 1998, p. 34.

⁷¹ Réponse au questionnaire sur biblio-fr de M. Y-A. Lacroix, bibliothèque nationale du Québec.

⁷² “ L'enregistrement d'une image numérique est une représentation fondée sur l'image d'un document document original et nous le traitons en conséquence comme une reproduction ”. P. CONWAY. *Conversion of microfilm to digital imagery : a demonstration project. Performance report on the production conversion phase of project open book*. Yale university library, Newhaven, 1996, annexe “ Report of the cataloging digital reproductions task force ”, non paginée.

D'après ce même rapport, il convient toutefois de se servir des possibilités du format Marc pour signaler le document numérique. Divers champs de la notice bibliographique seront ainsi utilisés, comme le présente l'annexe n°6 :

- le champs 590 décrit les différentes versions de l'image numérique, en en donnant, pour l'identifier, sa résolution et sa destination.

- Le champs 533 décrit les caractéristiques plus précises de la reproduction : lieu et date de fabrication, poids, résolution, support de stockage...

- Le champs 538 apparaît à l'Opac et donne l'adresse de consultation du document numérique.

- Le champs 856, transparent à l'Opac, permettra de faire le lien au niveau du catalogue avec le document numérique.

Il reste cependant à voir si le format Marc utilisé en France, et en particulier à Cujas, permettra l'intégration de ces données. Se pose aussi plus largement la question des compétences des catalogueurs à s'adapter à ce nouveau support, qu'il soit, comme ici considéré comme une reproduction, ou faisant l'objet d'une notice particulière. Dans le cas de la bibliothèque Cujas, le manque de temps et de moyen fera sans doute que la question du catalogage des documents numérisés ne pourra être posée avant un certain temps.

5. La diffusion et ses problèmes.

a. Diffusion locale ou Internet.

Pour simplifier, on peut envisager deux modes de diffusion des données numériques résultant d'un programme de numérisation. La première consiste en une diffusion restreinte, soit purement locale dans la mesure où seuls les utilisateurs de tel poste de l'établissement pourront y avoir accès, soit étendu à un groupe restreint, par exemple les personnes fonctionnant sur un réseau interne, type intranet, sur lequel a été placée la base numérique. La mise sur un réseau ouvert à tous, comme Internet, constitue la deuxième solution.

Notre propos n'est pas ici de disserter sur l'importance qu'a pris le Web depuis quelques années, ni sur la faible part de documents francophones, situation qui inquiète les ministères de la culture et de l'enseignement supérieur, qui y sont présents, mais plutôt d'en déterminer les conséquences sur la base numérisée.

Dans le cas de Cujas, le choix d'une mise à disposition sur Internet a été fortement envisagée dès le début de la réflexion autour du projet de numérisation, et ce, pour plusieurs raisons. Tout d'abord, pour la bibliothèque, qui dépend pour l'informatique de Paris-I, mais dont une grande partie du lectorat dépend de Paris-II, et en particulier un groupe d'historiens du droit assidus à la réserve, la mise à disposition localement ou même sur le réseau dépendant de Paris-I aurait entraîné une limitation des personnes susceptibles de consulter cette base. De plus, consciente que nombre des lecteurs de la réserve ne résident pas à Paris, la bibliothèque a envisagé l'installation sur Internet très rapidement, pour des motifs de diffusion, mais aussi afin d'enrichir un site Web encore trop restreint pour un établissement de cette importance. On espère ainsi permettre aux lecteurs d'avoir accès aux données indépendamment des horaires d'ouverture de la bibliothèque.

Ce choix implique par ailleurs une réflexion plus poussée sur la présentation des documents numérisés, dans la mesure où un site Web constitue une vitrine non négligeable de l'établissement. Ce ne sont plus seulement les spécialistes déjà habitués de la réserve et de ses ouvrages qui sont visés mais plus largement des étudiants de tous cycles, juristes bien sûr mais pourquoi pas historiens, linguistes, ... Il a donc été envisagé d'enrichir la simple mise à disposition des documents par une présentation générale sur Internet de la base intellectuelle et, pourquoi pas, plus précise de chaque document, à condition que la bibliothèque puisse trouver les compétences et les moyens humains nécessaires.

Cette décision pose le problème des droits, particulièrement crucial et nouveau, portant sur ces documents mis en libre accès.

b. Les problèmes de droit.

La diffusion sur écran, que ce soit localement, ou, encore plus, sur Internet pose en effet de nombreux problèmes de droit d'auteur qui sont encore loin d'être résolus. Pour une bibliothèque se pose plus que jamais le problème de "sauvegarder le nécessaire équilibre entre le respect des titulaires du droit de la propriété littéraire et artistique et celui de l'institution dont la mission est d'assurer la diffusion la plus large des collections d'œuvres ainsi constituées"⁷³.

Plus précisément, parmi les droits d'auteurs, sont concernés le droit de reproduction et le droit de représentation car "la consultation par affichage sur écran procède du droit de représentation, mais également du droit de reproduction, défini dans l'article L 122-3 du code de la propriété intellectuelle comme la fixation matérielle de l'œuvre par tous les procédés qui permettent de la communiquer au public de manière indirecte"⁷⁴. Inutile de préciser que les problèmes déjà existants de mise à disposition du public de textes numérisés en local se trouvent amplifiés dès qu'on aborde la possibilité de la mise à disposition sur Internet !

Face à l'absence de législation spécifique sur ce point de droit précis, deux initiatives devraient permettre de clarifier la situation en France. Tout d'abord les diverses sociétés d'auteurs se sont rendu compte d'un manque de lisibilité de leur part en ce qui concerne la perception des droits. En conséquence, certaines d'entre elles, en particulier la Société des auteurs, compositeurs et éditeurs de musique (S.A.C.E.M.), la Société des auteurs compositeurs dramatiques (S.A.C.D.), la Société civile des auteurs multimédias (S.C.A.M.) et la Société des auteurs arts graphiques et plastiques (A.D.A.G.P.) envisagent la création de la société SESAME, chargée de gérer la perception des droits d'auteurs auprès des utilisateurs⁷⁵.

Enfin, la B.n.F. continue, dans le cadre de la mise à disposition de textes récents en mode caractère, à traiter avec les auteurs et les éditeurs. On peut espérer

⁷³ V. GAME. "La constitution d'un fonds d'images et d'ouvrages numérisés par la B.n.F.", dans *Bulletin d'informations de l'A.B.F.*, n°172, 1996, p. 91.

⁷⁴ D. ROUILLARD. "Possibilités et contraintes de l'utilisation du cédérom en bibliothèque", dans *Du CD-Rom à la numérisation. Développer les documents numériques en bibliothèque*, Villeurbanne, I.F.B., 1997, p. 101.

⁷⁵ D. ROUILLARD, *op. cit.*, p. 108.

que les négociations menées permettront d'aboutir à une législation fondée sur la jurisprudence qui, à défaut de loi, servirait de recours juridique. Pour les ouvrages imprimés, deux grands accords ont été ainsi signés entre l'ancienne bibliothèque de France et le syndicat national de l'édition, le 4 novembre 1991 et le 15 mars 1993, établissant la nécessité de l'autorisation préalable des auteurs avant numérisation, la rémunération éventuelle de ceux-ci et de leurs éditeurs, la mention du droit moral des auteurs sur écran. Par contre, si un ouvrage fait l'objet d'une numérisation pour des motifs de conservation, l'autorisation des auteurs n'est pas exigée⁷⁶. Ces mesures constituent un début mais ne résolvent pas tout : des difficultés surgissent régulièrement à ce sujet et continuent à causer de gros soucis aux bibliothèques.

Le cas de Cujas s'avère toutefois beaucoup plus simple dans la mesure où il s'agit pour l'instant de numériser des documents anciens tombés depuis longtemps dans le domaine public. Tous les problèmes de droit d'auteur exposés ci-dessus se trouvent annihilés. Il ne subsiste en réalité que le droit patrimonial de la bibliothèque sur les documents qu'elle conserve. Différentes solutions se présentent alors : utilisation de filigrane, comme ce qui s'est pratiqué à la bibliothèque Vaticane⁷⁷, marquage visible dans un coin de l'écran ou marquage invisible permettant de coder certaines informations dans le fichier image. Là encore, même si ces possibilités existent et sont appliquées, il faut rester conscient qu'elles ne constituent pas un obstacle pour un informaticien ! De plus, " en général, les solutions de protection des images ne sont mises en œuvre que pour les images de haute résolution permettant des publications " ⁷⁸. Sans sous-estimer la valeur des ouvrages de la réserve de Cujas, on peut déjà considérer que le risque de récupération de ces textes à des fins commerciales reste minime, moindre en tout cas que pour un fonds d'enluminures numérisés. Par ailleurs et en admettant cette éventualité, le taux de résolution choisi,

⁷⁶ V. GAME, *op. cit.*, p. 92.

⁷⁷ Y. SORDET. *Restauration et transfert de support à la bibliothèque Vaticane*. Mémoire d'études, ENSSIB, 1998, DCB6, n°33, p.77.

⁷⁸ L.-M. ALBIGES. " Concevoir et réaliser une banque d'images pour un public non spécialisé ", dans *Du CD-rom à la numérisation. Développer les documents numériques en bibliothèque*, *op. cit.*, p. 127.

volontairement peu élevé pour la mise à disposition sur Internet, entraîne une restitution inférieure aux besoins de l'imprimerie.

6. Évolutions et mutations.

a. Les enjeux de Cujas.

La bibliothèque Cujas attend beaucoup de son entreprise de numérisation. Les enjeux en apparaissent multiples. Les objectifs, qui ont transparu déjà au fil des choix techniques et intellectuels envisagés, peuvent être regroupés en trois grands thèmes : permettre une meilleure conservation des documents, assurer un meilleur service au lecteur de la réserve et créer une certaine publicité autour de celle-ci.

Il apparaît qu'une meilleure conservation des documents anciens reste l'objectif premier de ce type de numérisation. On a affaire ici à une consultation d'un petit nombre d'ouvrages au sein de la réserve. Ces derniers ont déjà subi les outrages des ans, les dommages de mauvaises conditions de rangement et n'ont fait l'objet d'une politique de préservation que récemment. A cela s'ajoutent les manipulations des lecteurs, souvent peu au fait des précautions élémentaires, comme l'utilisation des lutrins pour les grands formats. Il s'est donc avéré urgent de décharger cette pression sur les collections en proposant des supports de substitution de qualité, permettant de préserver les ouvrages originaux qui ne seraient désormais plus consultables, sauf besoin précis nécessitant le recours à l'original. En se plaçant dans cette seule optique, un support classique comme le microfilm aurait pu tout aussi bien convenir. La numérisation a cependant été retenue, nonobstant les problèmes de préservation que cette technique implique, en raison des autres objectifs de la bibliothèque.

Le meilleur service au lecteur de la réserve représente le deuxième enjeu de la bibliothèque pour son projet de numérisation. Nous avons constaté en effet la faiblesse des horaires d'ouverture de la réserve, qui devraient sans doute encore diminuer faute de personnel qualifié, et la difficulté d'accès plus générale. La numérisation devrait pallier cet inconvénient grâce à une plus large diffusion par le biais d'Internet, qui offrira au lecteur la possibilité d'une consultation sur place, mais

aussi à distance, que ce soit à son domicile ou à partir des centres de recherche, bibliothèques d'instituts et facultés, sans les difficultés de commande. L'offre simultanée à plusieurs lecteurs d'un corpus choisi parmi les ouvrages les plus consultés devraient aussi faciliter la vie des amateurs de la réserve, se disputant souvent le même livre dont il n'existe pas forcément de double. Enfin, la numérisation offre aussi la possibilité pour les lecteurs de reproduire tout ou partie de l'ouvrage, sous forme électronique ou papier, et d'y ajouter des annotations personnelles. A l'heure actuelle, les conditions de reproduction restent en effet fortement réglementées pour des raisons évidentes de préservation des ouvrages. On peut grossièrement distinguer deux cas : les livres postérieurs à 1810 peuvent faire l'objet de photocopies, en fonction de leur état physique examiné par le personnel de la réserve. Les documents antérieurs à 1810 sont interdits à la photocopie. Les lecteurs sont toutefois autorisés à en faire des photographies avec leur matériel ou à en demander la reproduction par microfiche après accord de la direction. Dans la pratique, le recours à cette solution reste peu exploité en raison des coûts, la bibliothèque ne disposant pas de photographes à demeure, au contraire de Sainte-Geneviève ou de la Bibliothèque interuniversitaire de médecine. Il n'existe à l'heure actuelle que deux livres ayant fait l'objet de ce traitement. L'impossibilité de reproduction handicape fortement le travail du lecteur, contraint de recopier de larges pans de droit. En permettant un tirage ou une reproduction électronique, la bibliothèque offre ainsi un nouveau mode de consultation, fort attendu de son public, et la possibilité de reconstituer à domicile une bibliothèque virtuelle d'histoire de droit.

L'aspect " publicitaire " a fait fortement pencher la balance en faveur de la numérisation. La réserve reste confidentielle, y compris au sein des lecteurs de Cujas, et seuls quelques professeurs d'histoire du droit, et leurs étudiants, l'exploitent réellement. La mise à disposition de documents sur le site Web en fera connaître plus largement l'existence et devrait attirer à la réserve, et en conséquence à la bibliothèque, un public conquis par ce que l'on pourrait appeler un " produit d'appel ". La

direction espère aussi plus pratiquement en tirer des arguments supplémentaires pour montrer l'intérêt de la réserve à son autorité de tutelle. La présentation de statistiques de consultation du site pourrait en constituer un élément. Le but de la numérisation à Cujas reste en effet un objectif hautement politique, qui explique à lui seul la préférence accordée à ce support par rapport aux microformes. Pour l'instant, aucun crédit, aucun poste n'est affecté à la réserve du fait du peu d'intérêt accordé aux fonds patrimoniaux des bibliothèques universitaires à quelques exceptions près. Le phénomène de mode autour du numérique, et notamment au sein du ministère de l'Enseignement supérieur, est ainsi utilisé pour attirer des crédits destinés à la mise en valeur de fonds anciens. Cette opération permet aussi de parler de la réserve au sein de l'autorité de tutelle et de demander la création d'un poste de catégorie A, au moins à mi-temps, sur cette réserve, condition *sine qua non*, pour une meilleure exploitation.

Beaucoup d'enjeux à court ou moyen terme prévalent donc à Cujas, avec notamment l'importance politique de la numérisation. A plus long terme, s'est toutefois posé le problème de la pérennité de la technique choisie, dont nous verrons par la suite qu'elle reste loin d'être évidente. La bibliothèque se prononce en faveur d'une résolution plutôt forte pour du texte ainsi que des techniques largement répandues pour assurer une conservation de ses données numériques. Elle se réserve aussi la possibilité, dans le cadre d'un futur plan, d'enrichir son corpus au fur et à mesure de ses moyens, en fonction de la réception faite au premier projet et sans pour autant devoir convertir ce qui avait déjà été fait.

b. La question de la pérennité.

Ainsi, pour tout ce qui concerne les documents anciens, la pérennité du support de substitution reste un problème crucial. La question se pose aujourd'hui de savoir dans quelle mesure la numérisation répond à cette attente, et le choix des Archives américaines de continuer leur politique de microfilmage semble le nier. Plus proche de nous, cette interrogation a été formulée au cours de la journée d'études de

l'Association des conservateurs de bibliothèques.⁷⁹ occasionnant une querelle des Anciens et des Modernes, comme on pourra la caractériser, opposant les tenants du microfilm comme support à long terme et les adeptes de la numérisation. S'il en est ressorti clairement que la question de la conservation des données numériques restait d'actualité aujourd'hui mais s'atténuerait sans doute à l'avenir, il a été fait surtout état de l'avantage des supports de l'information numérique, plus stables qu'un microfilm, très inflammable, mais aussi des problèmes de lecture à l'avenir.

La "commission of preservation and access", organisme nord-américain, se demande elle aussi si l'on peut considérer la numérisation comme un moyen de préservation. Un rapport récent de H. Weber et M. Dor⁸⁰ semble accorder sa préférence au microfilm pour la préservation, en reconnaissant les avantages d'accès de la numérisation : "it takes the view that in preservation, the enormous potential of digitization for access should be combined with the stability of microfilm for long-term storage [...]. A strategy for the introduction of digitization into preservation projects could be derived : microfilm has continuing priority as a recording and storage medium because of its quality and stability over time. As a medium for document delivery, the digital form, with its advantages of swift and remote access, in a quality depending on the intended use, should be employed"⁸¹.

Toujours selon ce rapport, le microfilm garde en effet de nombreux avantages sur son concurrent :

- il est assuré d'une plus grande pérennité dans la mesure où il ne repose pas sur une technologie évolutive pouvant entraîner l'obsolescence de son appareil de lecture.

⁷⁹ *Les bibliothèques face à la numérisation : enjeux et méthodes*. Journée d'études de l'A.C.B., Valenciennes, 6 novembre 1998.

⁸⁰ H. WEBER et M. DOR. (Page consultée le 10 octobre 1998). Digitization as a means of preservation ?, [en ligne]. Adresse URL : [Http://www.clir.org/cpa/reports/digpres.htm](http://www.clir.org/cpa/reports/digpres.htm)

⁸¹ " Il apparaît que, pour la préservation, l'énorme potentiel de la numérisation pour l'accessibilité doit être associé avec la stabilité du microfilm pour la conservation à long terme. [...]. Une stratégie pour l'introduction de la numérisation dans les projets de préservation doit être déclinée : le microfilm a toujours la priorité comme support de fixation et de stockage à cause de ses qualités et de sa stabilité dans le temps. Comme support de diffusion du document, la forme numérique, avec les avantages de son accès rapide et à distance, devra être employée, avec une qualité dépendant de l'utilisation recherchée. "

- Ses coûts de production et de duplication restent faibles.
- L'utilisation de focus lui permet de s'adapter à des tailles variées de documents.

En définitive, il s'avérerait tout à fait rentable à long terme de réaliser un microfilmage, permettant à l'avenir de procéder à plusieurs numérisations du même document sans le manipuler à nouveau et pour un coût plus réduit puisque le microfilmage donnera la possibilité d'automatiser la numérisation.

Ces arguments apparaissent aujourd'hui comme très écoutés dans le milieu des bibliothèques françaises, puisque la Bibliothèque nationale de France, et donc ses pôles associés, procèdent à cette double mesure de préservation et que la Direction du Livre et de la Lecture en a fait l'une de ses recommandations les plus fermes.

Si ces arguments favorisent le microfilmage au détriment des possibilités de préservation de la numérisation, on peut toutefois objecter que cette étape risque d'entraîner une moindre qualité de numérisation, et de fait rendre nécessaire une seconde opération. Le numérique peut offrir des garanties de longévité à conditions de respecter certaines précautions :

- **Respect des normes.** Celui-ci garantit une forte utilisation de ces techniques règlementées et donc à l'avenir l'invention de solution de transfert en cas d'obsolescence. Le choix de technologies "propriétaires" doit par contre être mûrement réfléchi.

- **Suivi de l'évolution technologique.** Il faut éviter de se retrouver en retard d'une étape.

- **Numérisation de bonne qualité.**

- **Respect du support physique.** Comme nous l'avons vu plus haut, même s'il semble plus résistant dans la mesure où il ne pose pas trop de contraintes permanentes de conservation, des transferts réguliers garantissent une meilleure conservation, à faible coût.

Dans le cas plus particulier de Cujas, on est en présence d'un fonds extrêmement intéressant et riche, comme les utilisateurs s'accordent à le dire,

rassemblant des ouvrages que l'on comparera avantageusement. Chaque ouvrage pris en particulier, en tout cas ceux qui feront l'objet de la numérisation qui nous concerne, n'ont pas une valeur bibliophilique pouvant les faire considérer comme des documents uniques. S'il faut veiller à leur conservation en raison des outils qu'ils constituent, un microfilmage ne semblait pas s'imposer et le choix d'une numérisation à 300 dpi, résolution relativement élevée pour du texte suffira sans doute à garantir sa longévité.

Conclusion.

Ce mémoire se proposait de montrer comment la numérisation permet de mettre en valeur un fonds patrimonial, en s'appuyant sur l'exemple de la réserve de la bibliothèque interuniversitaire Cujas.

On a vu que le contexte français se prêtait tout à fait favorablement à ce genre d'entreprise, où l'exemple d'opérations pilotées par la D.L.L. ou l'expérience de la B.n.F. pouvaient servir de modèles. L'étude plus approfondie de la bibliothèque Cujas, des problèmes de sa réserve comme des caractéristiques des ouvrages qu'elle renferme, a pu montrer les ambitions liées au projet de numérisation, plus à même de réunir l'objectif de préservation des documents comme de mise en valeur de la collection. Une simple campagne de microfilmage eût peut-être suffi, si la diminution de la pression sur les collections avait constitué le seul impératif de cette entreprise. Nous avons vu que si ce souci s'avérait réel, on espérait en outre de la mise en place d'une base de textes numérisés une certaine publicité autour de la réserve, fonds remarquable de l'aveu même des utilisateurs, et qui mérite une plus grande attention.

Suite à ce mémoire, on peut envisager deux conclusions : la première sur les difficultés à associer à cette heure numérisation et préservation de documents anciens, et en particulier auprès des professionnels des bibliothèques encore attachés au support argentique ; dans un second temps, nous nous attacherons à montrer que la numérisation et sa diffusion sur Internet constituera pour la bibliothèque Cujas un moyen de mise en valeur que l'on peut rendre extrêmement efficace par quelques démarches complémentaires.

Tout d'abord, si la numérisation est dans le cas de Cujas apparu comme le moyen de mise en valeur le plus adapté, la " publicité " qui en découlera a un prix, celui de la technologie, puisqu'il s'agit de choisir une technique alliant performance et

respect du document et s'adaptant aux caractéristiques du livre ancien. L'enjeu s'avère d'autant plus important, le droit à l'erreur est d'autant plus réduit, qu'un aspect de préservation des documents numériques se greffe, pour éviter toute nouvelle manipulation à l'avenir : " chaque numérisation supplémentaire génère de nouveaux frais, sans oublier la possibilité d'endommager l'original par des manipulations excessives. Aujourd'hui déjà, le temps de travail (recherche des documents, nettoyage, numérisation proprement dite, remise en place de l'original) constitue l'élément le plus coûteux de l'opération ”⁸².

Pour cette raison, on ne peut en effet envisager de numériser à nouveau dans un temps rapproché, à la fois pour ces problèmes de coût évoqués, mais aussi de préservation des ouvrages : il convient de mener une opération où la question de la pérennité des données numériques sera résolue au mieux. " Cette sauvegarde numérique évite d'avoir sans cesse recours aux originaux, dont la durée de vie est ainsi prolongée et assure en même temps la conservation des données dans leur état actuel. Peut-on envisager une meilleure préservation préventive de notre patrimoine ? ”⁸³. Encore pour s'assurer de cette longévité du support de substitution faut-il accepter de rester à l'écoute des techniques, savoir s'y adapter en permanence, remettre en question et revoir périodiquement le concept d'archivage face à une accumulation de données.

Par ailleurs, second axe de notre conclusion, il convient de souligner que le choix d'Internet comme facteur de diffusion de la base numérisée devrait réellement répondre aux objectifs de mise en valeur recherchés par l'établissement pour sa réserve. Cette solution présente quelques inconvénients bien connus : une adaptation intellectuelle du produit devra être réalisée, par exemple une présentation des textes, mais aussi une adaptation physique par le transfert des données sur le serveur

⁸² R. GSCHWIND et A. GÜNZL. " Numériser pour conserver. Une solution pour l'archivage à long terme des collections photographiques ", dans *Informatique et conservation-restauration du patrimoine culturel. 8^e journées d'études de la S.F.I.I.C. Châlons-sur-Saône, 23-24 octobre 1997.* S.F.I.I.C., Champs-sur-Marne, 1997, p.134.

⁸³ R. GSCHWIND et A. GÜNZL, *op. cit.*, p. 131.

Internet de Cujas, dont il faudra s'assurer préalablement de la capacité en mémoire pour accueillir cette base importante. L'alliance Internet-numérisation des textes anciens devrait offrir à Cujas l'outil publicitaire et de diffusion dont elle manque cruellement pour sa réserve.

Toutefois, la diffusion sur Internet de cette collection numérique de textes juridiques anciens ne permettra une mise en valeur efficace de la réserve que si elle est relayée par une véritable opération de marketing. Pourquoi en effet ne pas considérer la numérisation comme un produit commercial destiné à créer un besoin de la réserve parmi les utilisateurs ? Le but avoué de cette opération reste de faire une publicité autour de la réserve, et partant, il convient d'appliquer au monde des bibliothèques des techniques adaptées du monde commercial.

On pourra donc, pour faire de cette numérisation à Cujas un véritable instrument de mise en valeur de la réserve, l'accompagner de différentes réalisations destinées à en faciliter l'accès : tout d'abord, il faudra susciter l'intérêt, voire une demande, de la part des utilisateurs autour de ce produit numérique, en les informant préalablement, en fabricant des plaquettes explicatives, en réalisant une rapide enquête auprès des lecteurs pour évaluer les demandes précises, en écrivant quelques articles autour de ce projet dans les revues professionnelles ou dans les revues étudiantes des facultés concernées et enfin en présentant la base dans les périodiques spécialisés en droit.

Deuxième idée pour améliorer l'impact de cette numérisation, développer des outils pour permettre au lecteur d'apprécier cette base à sa juste valeur : effectuer des démonstrations, notamment auprès des professeurs prescripteurs de consultation, inclure au sein du programme de formation aux utilisateurs quelques exercices pratiques d'orientation autour de cette collection numérique. Il conviendra aussi de former le personnel de la bibliothèque à renvoyer le lecteur vers cette base immatérielle.

On peut aussi espérer beaucoup de la mise en place sur Internet, à côté de cette base numérique de " travail " et réservée de prime abord au lecteur confirmé, une

base plus ludique destinée à attirer un plus large public, sous la forme par exemple d'une exposition virtuelle à l'image de ce que réalise actuellement la Bibliothèque interuniversitaire de Médecine.

Enfin, il faut offrir à l'utilisateur de la collection numérique les outils avec lesquels il pourra s'approprier un mode nouveau de consultation de documents, par exemple la fonction "stabilo" permettant de surligner les passages intéressants.

En définitive, une numérisation bien choisie et conçue en fonction des objectifs de la bibliothèque peut effectivement s'avérer un excellent mode de mise en valeur de fonds patrimoniaux, à condition d'en faire un véritable outil interactif, pont entre le lecteur et le bibliothécaire. Il faudra surtout éviter de faire de cette technologie performante un nouveau support de substitution, certes plus performant sur le plan de la restitution de l'image que le microfilm, mais tout aussi rebutant que ce dernier pour le lecteur. C'est le risque que l'on peut courir à ne pas encadrer le lecteur dans sa découverte de cette nouvelle technologie et des nouveaux modes de travail qu'elle implique.

Annexes

1. Glossaire des termes techniques cités.

BIT : Contraction de Binary Digit. L'un des chiffres 0 et 1, lorsqu'il est employé en numération binaire.

BITONAL : Technique de numérisation consistant à coder l'information sur un bit, c'est-à-dire que les valeurs 0 et 1 représentant chacune une couleur : noir ou blanc.

BITONAL TRAME : Technique de numérisation découlant de la précédente mais à laquelle on ajoute pour chaque image un tramage donnant l'impression d'une numérisation en niveau de gris.

C.C.I.T. : Abréviation de Comité Consultatif International de Téléphonie, à l'origine de différentes normes de compression de fichiers pour les documents monochromes. Le C.C.I.T.-groupe IV est la norme concernant les fax mais est aussi utilisée pour la compression de documents numériques en bitonal.

CD-ROM : Abréviation de Compact-Disk Read only memory. Disque optique normalisé de 12 cm et supportant 650 Mo de mémoire. Support non réinscriptible et non effaçable.

COMPRESSION : Opération destinée à favoriser le stockage et la transmission des données à moindre coût. Elle consiste à réduire le volume des données. La même partie du logiciel de compression permettra la restitution de l'intégralité de l'information de base en mode texte ou une qualité d'image suffisante pour être exploitée.

DONNÉE : Représentation d'une information élémentaire dans un format qui permet de lui appliquer un traitement.

D.P.I. : Abréviation de Dots per Inch. Unité de mesure de la résolution.

DVD-ROM : Abréviation de Digital Versatil Disk-Read Only Memory. Disque optique de 12 cm et contenant de 4,7 à 18 Go de mémoire. Support non réinscriptible.

FORMAT : Structure qui caractérise la présentation des informations au sein d'un ordinateur lors d'une transmission sur un support d'entrée. Le format d'édition est le format qui permet de préciser la mise en page des résultats sur le document final. T.I.F.F. en est un exemple.

G.I.F. : Abréviaton de Graphic Interchange Format. Suffixe indiquant qu'un fichier contient une image graphique. Format de fichier massivement utilisé pour compresser les images des pages Web sur Internet. Il a été développé à l'origine par la société américaine CompuServe.

GRAPHIC INTERCHANGE FORMAT *Voir* G.I.F.

H.T.M.L. : Abréviaton de Hypertext mark-up langage. Langage de balisage qui se présente comme une implémentation simplifiée de S.G.M.L. Il est particulièrement répandu sur Internet.

INTERNET : Interconnexion d'un grand nombre de réseaux informatiques mondiaux, nationaux et régionaux qui ont choisi d'utiliser le même protocole de communication, TCP/IP (Internet protocole) qui permet à des paquets de données de traverser les multiples réseaux pour arriver à la machine destinataire.

J.B.I.G. : Abréviaton de Joint Bilevel Expert Group. Standard de compression et de stockage d'images à deux niveaux (noir et blanc). Défini dans la norme ISO-11544.

JOINT PHOTOGRAPHIC EXPERTS GROUP *Voir* J.P.E.G.

J.P.E.G. : Acronyme de Joint Photographic Experts Group. Norme répandue de numérisation pour compresser fortement des images fixes au prix d'une certaine dégradation. C'est un format souvent utilisé sur Internet. Défini dans la norme ISO-10918.

MODE TEXTE : Technique de numérisation aboutissant à l'obtention d'un document électronique de type textuel permettant ainsi une recherche plein texte.

MODE IMAGE : Technique de numérisation aboutissant à l'obtention d'un fac-similé électronique du document original.

NIVEAU DE GRIS : Technique de numérisation consistant à coder l'information sur plusieurs bits de façon à élargir la gamme des gris qui permettront de restituer l'image initiale.

NIVEAU DE COULEUR : Technique de numérisation consistant à restituer une image en la décomposant en trois couleurs, rouge, vert, bleu (système RVB) dont chacune sera obtenue par codage sur un ou plusieurs bits selon l'éventail de couleurs souhaité.

NUMÉRISATION : Procédé utilisé pour transférer un texte, une photo, de la vidéo, du son sur un support informatique. Pour ce faire, ces informations sont transformées en données binaires afin d'être comprises, manipulées et transmises par un ordinateur. La numérisation d'une photo, par exemple, peut se réaliser au moyen d'un scanner, d'une capture vidéo (à partir d'une caméra ou d'un magnétoscope) ou d'une image de synthèse.

O.C.R. : Abréviation de Optique Character Recognition. Terme anglosaxon pour R.O.C.

P.D.F. : Abréviation de Portable Document Format. Format de documents portables, développé par Adobe. Il permet la consultation uniquement de documents sur diverses plates-formes.

PIXEL : Ou élément d'image. Élément bidimensionnel le plus petit d'une image auquel on puisse affecter des attributs tels que couleur et intensité.

RECONNAISSANCE OPTIQUE DE CARACTÈRES *Voir R.O.C.*

RÉSOLUTION : Nombre de points par pouce carré d'une image. Plus le nombre de points est élevé, plus la qualité de l'image est bonne. Elle s'exprime en d.p.i.

R.O.C. : Acronyme de Reconnaissance optique de caractères. Technique qui permet de transformer, via un scanner et un logiciel adéquat, les textes manuscrits ou imprimés (constituant un seul bloc non modifiable) en fichier texte de caractères indépendants pouvant donc être modifiés à volonté.

SCANNER : Numériseur. Appareil transformant la luminosité et les couleurs d'un document imprimé (texte, photo) en données numériques. Ces données peuvent alors être stockées, retraitées, imprimées,...

T.I.F.F. : Acronyme de Tag Image File Format. Format souvent utilisé pour les fichiers générés par les scanners. Il existe plusieurs possibilités suivant le nombre de couleurs choisies. Ce format n'engendre pas de perte mais ne permet qu'une faible compression.

U.I.T. : Abréviation de Union internationale des Télécommunications. Nouveau nom pour C.C.I.T. La norme de compression U.I.T.-IV correspond donc à l'ancienne C.C.I.T.-IV.

WORLD WIDE WEB : Application la plus célèbre d'Internet.

2. Photocopies de pages significatives des ouvrages de la Réserve.

A. Azon, *Summa Azonis...*, 8° ; cote 40629 ; p. de titre. *Écriture bicolore rouge et noire avec encadrement ; exemple de gravure sur bois pleine page avec écriture manuscrite.*

B. Azon, *Summa Azonis...*, 8° ; cote 40629 ; fol. CCXIII. *Écriture gothique sur deux colonnes avec gloses en regard. Papier faiblement et uniformément jauni.*

C. Burlamaqui J.J., *Principes de droit politique...*, 4° ; cote 13338 ; p. 319. *Exemple de page 4° saine.*

D. Coras, J. de, *Opera... omnia*, fol. ; cote 1067 ; page signée):(. *Taches brunes de roussissures réparties dans le texte et la marge. Exemple de page non numérotée, avec différentes typographies.*

E. Diderot D. et D'Alembert J., *Encyclopédie*, 4° ; cote 10.513.7, p. 65. *Exemple de forte tache d'humidité.*

F. Bodin J., *Les six livres de la République*, fol. ; cote 1318 ; page de titre. *Page de titre portant des traces de moisissures anciennes et différentes écritures manuscrites.*

___ G. Bodin J., *Les six livres de la République*, fol. ; cote 1318 ; p. 103. *Page portant une faible trace d'humidité. Différentes typographies avec abréviations et ligatures.*

3. Exemples de numérisation de quelques pages en bitonal, niveau de gris et niveau de couleur.

Nous avons présenté, pour chaque page proposée et décrite ci-dessous, trois résultats de numérisation suivant trois techniques avec dans l'ordre : bitonal, niveau de gris et niveau de couleur.

1. Jousse, D., *Nouveau commentaire sur l'ordonnance civile du mois d'avril 1667*, 12° ; cote 40718 ; p. XII. *Page saine.*
2. Capon, J., *Controversarium forensium*, fol. ; cote 146 ; portrait de l'auteur. *Exemple de gravure sur cuivre pleine page.*
3. Capon, J., *Controversarium forensium*, fol. ; cote 146 ; p. 36. *Page parsemée de taches brunes de roussissures.*
4. Diderot D. et D'Alembert J., *Encyclopédie*, 4° ; cote 10.513.7, p. 105. *Forte trace d'humidité.*
5. Diderot D. et D'Alembert J., *Encyclopédie*, 4° ; cote 10.513.7, p. 264. *Faible trace d'humidité.*
6. Vincentinus Jordanus, *Elucubrationem diversarum*, fol., cote 144 ; p. 458. *Espace typographique ayant viré au brun quasi-uniformément.*

4. Grille préparatoire appliquée au document avant numérisation.

I. Identification de l'ouvrage.

Cote
Auteur
Titre
Format
Nombre de volumes

II. État physique de l'ouvrage.

- *Reliure*

Plat supérieur
Plat inférieur
Dos
Couture
État de conservation des nerfs
Tranchefiles
Souplesse d'ouverture permise par la reliure

- *Corps de l'ouvrage*

état des feuilles
- pages déchirées avec pertes d'information : nombre et localisation
- pages déchirées sans perte d'information : nombre et localisation
- pages froissées/pliées
- pages dépassant le format habituel
état des fonds de cahier
état des illustrations
état du papier (moisissures)

III. Signalisation des anomalies.

- *Pages*

existence de pages manquantes : nombre et localisation
Inversion de cahiers
Taches d'humidités
Degré de roussissure
Nombre de pages à numériser
Pages à numériser (dans le cas d'un ouvrage numérisé partiellement)

- *Foliotation*

Anomalie de foliotation
Foliotation discontinue (ex. Romaine/arabe)
Existence de pages non numérotées : nombre et emplacement

Foliotation ou pagination

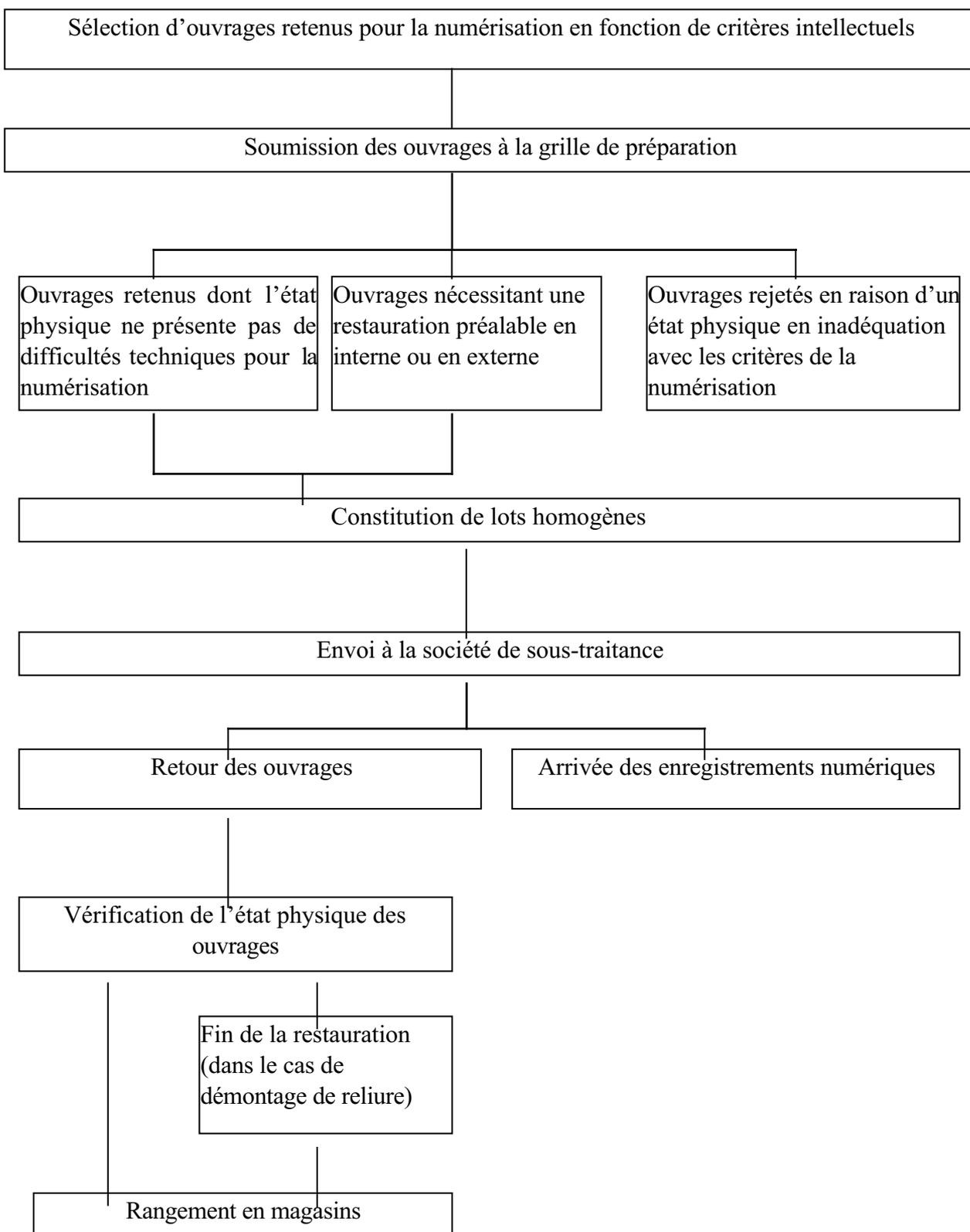
- *Table des matières*
 - Nombre de pages ou de lignes
 - Langue
 - Typographies (abréviation à restituer)
 - Pages précises

- *Annotations*
 - Localisation
 - Nombre
 - Type d'encre
 - Taille du caractère

- *Dessins et couleurs*
 - Existence d'illustrations
 - Localisation
 - Type
 - Nombre
 - Format : pleine page/dépl.
 - Page de titre bicolore
 - Rubrication

- *Numérisation supplémentaire.*
 - Nombre
 - Localisation
 - Support

5. Circuit interne du document pour sa numérisation.



6. Exemple de notice d'un document numérisé.

7. La communication à Cujas en 1997. Chiffres issus du rapport d'activité de 1997.

Lecteurs inscrits.

Monde universitaire	20830
- 1 ^{er} cycle	1287
- 2 ^e cycle	6442
- 3 ^e cycle	8673
- doctorants/chercheurs	3110
- ATER	739
- Professeurs	579
Autres (fonctionnaires, professionnels, ...)	4464
Total	25294

Nombre de places disponibles.

- Ensemble de Cujas	650
- Réserve Pavillon	7

Nombre d'heures d'ouverture.

- Ensemble de Cujas	2770 h.
---------------------	---------

Nombre de volumes.

- Ensemble de Cujas	environ 600 000
- Réserve	environ 10 000

Nombre de demandes de communication.

- Ensemble de Cujas	244 265 demandes
- salle de lecture générale	169413 demandes
- réserve	358 demandes dont 14% non consultées.

8. Tableau récapitulatif Mode texte/Mode image.

Mode image	
Avantages	<ul style="list-style-type: none"> • Toujours réalisable • Restitution des caractéristiques de la page • Coût plus faible
Inconvénients	<ul style="list-style-type: none"> • Volume en mémoire jusqu'à 8 fois plus élevé • Pas de possibilité de recherche textuelle

Mode texte	
Avantages	<ul style="list-style-type: none"> • Facilité de compression \Rightarrow faible volume en mémoire • Possibilité de recherche textuelle • Possibilité de révision et de remise en forme
Inconvénients	<ul style="list-style-type: none"> • Pas de restitution de la page initiale • Outil non fiable à 100% dans le cas de l'utilisation d'un O.C.R. <ul style="list-style-type: none"> <i>Obstacles à fiabilité :</i> <ul style="list-style-type: none"> ◇ Résolution du scanner ◇ Génération du document électronique ◇ Reliure de l'original ◇ Complexité linguistique <i>Obstacles à la fiabilité plus caractéristique au livre ancien :</i> <ul style="list-style-type: none"> ◇ Typographie anciennes (lettres couplées, caractères gothiques,...) ◇ Papier jauni entraînant un mauvais contraste ◇ Papier taché d'humidité ou de roussissure entraînant une difficulté de distinction du caractère • Nécessité d'une vérification automatique et humaine • Coût en résultant

9. Tableau récapitulatif sur les supports de stockage⁸⁴.

Supports magnétiques

Produits	Date de création et description	Utilisation	Lecteur
Bande magnétique classique	support d'informations à accès séquentiel	Sauvegarde	Lecteur spécifique
Digital Audio Tape (DAT)	Bande audionumérique	Sons, textes, images	Lecteur spécifique
Disque magnétique amovible	Cartouche de stockage jusqu'à 135 Mo (ex. : disque Syquest)	Sons, textes, images	Lecteur spécifique

Supports optiques

Produits	Date de création et description	Utilisation	Lecteur
Disque compact informatique (cédérom)	1985 Seule la lecture est possible	Images fixes, sons. Peu adapté au multimédia.	Lecteur de CD-ROM
Disque inscriptible (CD-R ou CD-WORM)	Disque inscriptible en monosession ou multisession. Surface dorée	Textes, sons, images	Du type d'enregistrement dépend le lecteur.
DON WORM	Disque de 30 (10 à 16 Go de capacité) ou 35.5 (25 Go) cm destiné à l'archivage de masse. Pas de standard : variété des procédés d'écriture. Enregistrement permanent.	Textes, sons, images à conserver sur une longue durée (plusieurs dizaines d'années)	Enregistreur/lecteur spécifique
DVD-R (Disque universel numérique)	1997 12 cm ; capacité de 7,9 Go	Inscriptible mais non effaçable ; sauvegarde de	Lecteur-enregistreur spécifique

⁸⁴ Extrait de : G. BEQUET. (Page consultée le 9 novembre 1998). La numérisation des documents patrimoniaux, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/mrt/conservation/fr/preventi/guide_dll.htm.

enregistrable)		données non compressées	
-----------------------	--	----------------------------	--

10. Tableau récapitulatif microfilm/numérisation.

	Microfilm	Numérique
Avantages	<ul style="list-style-type: none"> • Coût de production et de duplication • Pérennité car ne reposant pas sur une technique évolutive • Focus permettant une adaptation toute taille de document 	<ul style="list-style-type: none"> • Données binaires non dégradées en fonction du temps et de la génération du document • Facilité de diffusion et d'accès • Facilité de lecture • Coût de duplication • Qualité de restitution • Possibilité d'intervention sur le document
Inconvénients	<ul style="list-style-type: none"> • Support fragile (feu) • Dégradation des données avec le temps • Dégradation des données entre le document primaire et secondaire au moment de la duplication • Pas de possibilité d'intervention sur le document obtenu • Difficulté de lecture 	<ul style="list-style-type: none"> • Coût de production • Dépendance par rapport à des formats de fichier, des logiciels de traitement et du matériel de lecture en modification permanente • Peu au point sur des documents de grande taille
Points communs	<ul style="list-style-type: none"> • Mêmes contraintes physiques sur le document au moment de la prise de vue 	

Bibliographie.

1. Bibliographie.

BOUCHER (Thierry). *La numérisation des documents imprimés à la B.n.F.*, Enssib, 1994, mémoire DCB, 1994, n°28, 71 ff.

BEQUET (Gaëlle). “ Bibliotheca universalis : vers un catalogue commun des documents numérisés accessibles par les réseaux télématiques ”, *Bulletin des Bibliothèques de France*, 1995, t. 40, p. 54-57.

(—). “ La numérisation des collections patrimoniales : les enjeux ”, intervention du 17 juin 1998 au cours du séminaire organisé par l’E.N.S.S.I.B. *Conservation et mise en valeur des collections patrimoniales*.

CONWAY (Paul). *Conversion of microfilm to digital imagery : a demonstration project. Performance report on the production conversion phase of project open book*. Yale university library : New Heaven, 1996.

COCQ (Dominique). “ Le rôle des institutions en matière de conservation et de mise en valeur du patrimoine écrit : la politique de la Direction du Livre et de la Lecture hier et aujourd’hui ”, intervention du 16 juin 1998 au cours du séminaire organisé par l’E.N.S.S.I.B. *Conservation et mise en valeur des collections patrimoniales*.

DUCHARME (Christian). *Du CD-ROM à la numérisation : développer les documents numériques en bibliothèque*. I.F.B. : Villeurbanne, 1997. 172 p. (Boîte à outils, 5).

FROSSARD (F.). “ Stockage amovible : à chaque technologie son mode d’utilisation ”, *01 Informatique*, n°1481, 16 janvier 1998, p. 34.

GAME (Valérie). “ La constitution d’un fonds d’images numériques et d’ouvrages numérisés par la B.n.F. ”, *Bulletin d’informations de l’Association des bibliothécaires de France*, 1996, n°172, p. 89-92.

GLADNEY (H.M.), MINTZER (F.) et SCHIATARELLA (F.). “ Digital access to antiquities ”, *Communication of the A.C.M.*, vol.4, n°4, 1998, p. 49-57.

HERMEL (L.) et NICOLAS (Anne-Laure). *Glossaire multimédia et Internet*. Paris : Economica, 1997. 109 p.

Informatique et conservation-restauration du patrimoine culturel. 8^e journées d'études de la S.F.I.I.C. Châlons-sur-Saône, 23-24 octobre 1997. S.F.I.I.C. : Champs-sur-Marne, 1997. 303 p.

LEYRIS (Jean-Claude). “Mémoire d'un patrimoine”, *Bulletin des bibliothèques de France*, t. 42, n°3, 1997, p.

MAIGNIEN (Yannick). “La constitution de la collection numérisée de la B.n.F. : vers un nouvel encyclopédisme”, *Literary and linguistic computing*, vol. 10, n°1, 1995, p. 51.

MAIGNIEN (Yannick) et WAGNEUR (Jean-Didier). “Numérisation et nouvelles pratiques de lecture”, *Bulletin d'informations de l'Association des bibliothécaires de France*, 1995, n°167, p. 39-42.

Patrimoine et multimédias : le rôle du conservateur. Actes du colloque organisé par l'École du patrimoine, tenu à la B.n.F. les 23, 24 et 25 octobre 1996 Paris : Documentation française-École nationale du Patrimoine, 1997. 331 p.

RICHARD (Michel). “Le programme numérisation de la bibliothèque de France”, *Bulletin des bibliothèques de France*, 1993, n°3, p. 53-63.

SORDET (Yann). *Restauration et transferts de support à la bibliothèque vaticane*, mémoire d'études, EN.S.S.I.B., 1998, DCB 6, n°33. 110 ff.

ZOLTAN (Nemes). *Conception d'un système documentaire avec numérisation de documents pour la documentation régionale de la bibliothèque municipale de la Part-Dieu*. Rapport de stage, E.N.S.S.I.B.-Lyon 1, 1992, D.E.S.S.I.D., 62 ff.

2. Sites.

BEQUET (Gaelle). (Page consultée le 9 novembre 1998). La numérisation des documents patrimoniaux, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/conservation/fr/preventi/guide_dll.htm.

(—), B. CARBONE, P. DEBRIEN et al. (Page consultée le 5 mai 1998). Chapitre 8. La bibliothèque multimédia, [en ligne]. Adresse URL : <http://www.culture.fr/culture/mrt/numerisation/fr/dll.htm>.

DALBERA (Jean-Pierre). (Page consultée le 5 mai 1998). Des bases de données à l'Internet : 25 ans d'innovations au service de la culture, [en ligne]. Adresse URL : <http://www.culture.gouv.fr/culture/mrt/numerisation/fr/politique.htm>.

(—). (Page consultée le 9 novembre 1998). Patrimoine culturel et société de l'information, [en ligne]. Adresse URL : <http://www.culture.gouv.fr/culture/mrt/numerisation/intro-generale.htm>.

- DEVILLE (Jacques). (Page consultée le 9 novembre 1998). La reproduction des documents patrimoniaux. Problématique générale, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/conservation/fr/preventi/guide_dll.htm.
- FLEISCHAUER (C.). (Page consultée le 30 septembre 1998). Digital formats for content reproductions, [en ligne]. Adresse URL : <http://memory.loc.gov/ammem/formats.htm>.
- FRANCE. MINISTÈRE DE LA CULTURE ET DE LA COMMUNICATION. (Page consultée le 5 mai 1998). Appel à projets pour la numérisation de la documentation sur le patrimoine culturel en région, [en ligne]. Adresse URL : http://www.culture.gouv.fr/culture/mrt/numerisation/numerisation_programmes.htm
- HAIGH (Susan). (Page consultée le 9 novembre 1998). Les ressources dans les bibliothèques canadiennes : analyse d'une enquête de la bibliothèque nationale ", [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/cidl/fsurvey.htm>.
- (—).(Page consultée le 9 juillet 1998). Le reconnaissance optique des caractères (R.O.C.) en tant que technologie de numérisation, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/pubs/netnotes/fnotes37.htm>.
- KNOLL (A.). (Page consultée le 9 juillet 1998). Digitization of old manuscripts in the National Library of the Czech Republic in Prague, [en ligne]. Adresse URL : <Http://www.nkp.cz/digit/digit.htm>.
- KUNY (T.). (Page consultée le 8 juillet 1998). Introduction aux technologies et aux problèmes de la numérisation, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/pubs/netnotes/fnotes14.htm>.
- TEASDALE (Guy). (Page consultée le 30 septembre 1998). Le projet Gutenberg, [en ligne]. Adresse URL : http://www.bibl.ulaval.ca/info/pagepers/teasdale/LBQ/lettre_12.htm.
- UNIVERSITÉ LAVAL. BIBLIOTHÈQUE. (Page consultée le 30 septembre 1998). Notre mémoire en ligne. Description détaillée du projet, [en ligne]. Adresse URL : <http://www.nlc-bnc.ca/cihm/ecol/francais/desc-fr.htm>.
- WEBER (Harmut) et DOR (Mariane). (Page consultée le 10 octobre 1998). Digitization as a method of preservation ?, [en ligne]. Adresse URL : <Http://www.clir.org/cpa/reports/digpres.htm>.

Table des matières.

<u>INTRODUCTION</u>	1
<u>I. LA BIBLIOTHÈQUE INTERUNIVERSITAIRE CUJAS</u>	7
1. PRÉSENTATION RAPIDE DE LA BIBLIOTHÈQUE CUJAS.....	7
2. PRÉSENTATION DE LA RÉSERVE DE LA BIBLIOTHÈQUE CUJAS.....	10
<i>a. La réserve et sa mise en place</i>	10
Historique de la formation de la bibliothèque.....	10
Constitution de la réserve.....	12
<i>b. Intérêt du fonds ancien</i>	14
Composition générale du fonds.....	14
Particularité du fonds imprimé.....	14
<i>c. Les usages des lecteurs</i>	17
<i>d. Les difficultés rencontrées pour la mise en valeur de la réserve</i>	19
<u>II. LE CONTEXTE GÉNÉRAL</u>	23
1. DÉFINITION DE LA NUMÉRISATION.....	23
2. LES ENJEUX DE LA NUMÉRISATION.....	26
3. LE CADRE FRANÇAIS : LES MOYENS OFFERTS PAR L'ÉTAT.....	27
4. PRÉSENTATION DE QUELQUES EXEMPLES DE NUMÉRISATION EN BIBLIOTHÈQUE.....	31
<i>a. La bibliothèque nationale de Tchéquie : la numérisation de documents précieux</i>	32
<i>b. La Bibliothèque nationale du Canada : le choix de retracer l'histoire d'un pays</i>	33
<i>c. Un projet à visée encyclopédique : la Bibliothèque nationale de France</i>	35
<u>III. RÉFLEXION AUTOUR D'UNE NUMÉRISATION DE LIVRES ANCIENS À TRAVERS L'EXEMPLE DE CUJAS</u>	37
1. LA PRÉPARATION DES DOCUMENTS.....	37

2. LA NUMÉRISATION.	40
<i>a. Choix du matériel pour numériser</i>	40
<i>b. Le choix des paramètres de numérisation</i>	41
Mode image ou mode texte ?.....	41
Choix du mode de traitement.....	43
Choix de la résolution.....	48
<i>c. Le contrôle qualité</i>	48
3. LES OUTILS D'ACCÈS.	50
<i>a. La création de la table de correspondance</i>	50
<i>b. L'accès par la notice</i>	53
<i>c. L'accès par la table des matières</i>	55
4. LE STOCKAGE.....	57
<i>a. Support de stockage</i>	57
<i>b. Support de diffusion</i>	58
<i>c. Catalogage des données numériques</i>	59
5. LA DIFFUSION ET SES PROBLÈMES.....	60
<i>a. Diffusion locale ou Internet</i>	60
<i>b. Les problèmes de droit</i>	62
6. ÉVOLUTIONS ET MUTATIONS.	64
<i>a. Les enjeux de Cujas</i>	64
<i>b. La question de la pérennité</i>	66
<u>CONCLUSION</u>	70
ANNEXES.....	I
1. <u>GLOSSAIRE DES TERMES TECHNIQUES CITÉS</u>	II
2. <u>PHOTOCOPIES DE PAGES SIGNIFICATIVES DES OUVRAGES DE LA RÉSERVE</u>	VI
3. <u>EXEMPLES DE NUMÉRISATION DE QUELQUES PAGES EN BITONAL, NIVEAU DE GRIS ET NIVEAU DE COULEUR</u>	VII
4. <u>GRILLE PRÉPARATOIRE APPLIQUÉE AU DOCUMENT AVANT NUMÉRISATION</u>	VIII

<u>5. CIRCUIT INTERNE DU DOCUMENT POUR SA NUMÉRISATION.....</u>	X
<u>6. EXEMPLE DE NOTICE D'UN DOCUMENT NUMÉRISÉ.</u>	XI
<u>7. LA COMMUNICATION À CUJAS EN 1997.CHIFFRES ISSUS DU RAPPORT D'ACTIVITÉ DE 1997.</u>	XII
<u>8. TABLEAU RÉCAPITULATIF MODE TEXTE/MODE IMAGE.....</u>	XIII
<u>9. TABLEAU RÉCAPITULATIF SUR LES SUPPORTS DE STOCKAGE.</u>	XIV
<u>10. TABLEAU RÉCAPITULATIF MICROFILM/NUMÉRISATION.</u>	XVI
BIBLIOGRAPHIE.	XVII
<u>1. BIBLIOGRAPHIE.</u>	XVII
<u>2. SITES.....</u>	XVIII