

Des puces, des livres... et des hommes

La numérisation des imprimés à la Bibliothèque nationale de France °

Julie Bouchard

Dès la fin de l'été 1998¹, la Bibliothèque nationale de France (BNF) offrira à ses lecteurs universitaires une ambitieuse collection de 100 000 documents "numérisés"². Ceux-ci seront consultables sur des postes de lecture assistée par ordinateur (PLAO), développés par la Bibliothèque elle-même à l'intention des lecteurs savants. En principe, ils seront aussi diffusables à distance: les initiateurs du projet semblent avoir anticipé le succès public d'Internet.

Cependant, le programme de numérisation des imprimés de la BNF ne se résume pas à une somme d'octets qui attendent au fond de leur serveur qu'on les sorte pour être lus. Il est également le résultat d'une série de choix tant techniques, qu'économiques ou culturels. Conçu et lancé sur le ton de l'optimisme et du volontarisme d'État, le programme de numérisation a été soumis à un fort "*ancrage social*"³, qui l'a petit à petit éloigné de ses ambitions initiales.

L'objectif central était la transmission à distance des documents; au *finish*, la collection numérisée risque de rester prisonnière des murs de la Bibliothèque parisienne et des fameux PLAO. Pourquoi et comment est survenu ce "détournement" majeur? L'ambition du projet élaboré dans une pure logique d'offre semble victime d'elle-même, ayant fait l'impasse sur un ensemble de considérations socio-techniques. Huit ans après ses débuts, le programme de numérisation relève toujours du pari et les incertitudes sont à la hauteur des ambitions: pharaoniques.

° Je remercie J.-J. Salomon (professeur), G. Schméder et C. Bertho-Lavenir (maîtres de conférences) du *Centre Science, Technologie et Société* du Conservatoire national des arts et métiers (CNAM) pour leurs conseils, critiques et encouragements, tant pour cet article que pour le mémoire de DEA sur lequel il est fondé. Mes remerciements vont aussi aux membres de la BNF qui ont eu la patience de répondre à mes questions, P. BÉLAVAL (Maître de requête au Conseil d'État, Directeur général), M.-N. DARMOIS (Chargée de mission à la Présidence), A. GIFFARD (Administrateur civil, ancien Directeur du Département de l'informatique et des nouvelles technologies), Y. MAIGNIEN (Responsable de la Mission scientifique de numérisation), A. BERNARD (Déléguée pour l'informatique à la Mission centrale de coordination et de planification), C. LAMARRE (Documentaliste au Centre de documentation), et M. DE FUENTES (Conservateur de la Bibliothèque de l'École polytechnique).

¹ Annonce faite par le Président de la BNF le 25 mars 1996 au Salon du Livre.

² La numérisation des imprimés est une sorte de photocopie informatique: cela consiste à convertir les informations imprimées sur le papier en une succession d'éléments binaires stockés sur supports numériques (disquettes et disques durs, CD-Rom, cassettes DAT, etc.). Numérisé, le livre devient communicable à distance via les réseaux de télécommunication et manipulable par un ordinateur.

³ L'expression est empruntée à Roger BAUTIER, *De la rhétorique à la communication*, Grenoble, Presses Universitaires de Grenoble, 1994, 319 p.

A. – DES MANUSCRITS AUX TECHNOLOGIES NUMÉRIQUES

Au XVII^e siècle, à l'époque de Colbert, la Bibliothèque du Roy sert une politique de prestige. Le ministre de Louis XIV déménage la Bibliothèque... près de son hôtel particulier. Puis il nomme des fidèles aux postes clés. Il fait acquérir de nombreux livres italiens et orientaux (alors très en vogue), des cartes, des ouvrages divers sur les civilisations lointaines, des estampes et des médailles. Dotée d'une collection des titres et des généalogies, la Bibliothèque joue alors à l'inspecteur des impôts: posséder les titres originaux permet l'identification des vrais et des faux nobles pour que soient attribués correctement les privilèges fiscaux réservés aux "vrais" aristocrates⁴!

Plus de trois siècles après cette royale époque, devenue propriété de la Nation en 1791 et désormais appelée Bibliothèque nationale, la Bibliothèque du Roy entre dans le XXI^e siècle. En 1988, un Président de la République, François Mitterrand, décide en effet d'en faire le bijou de son deuxième septennat. Maître à bord, la Présidence décide des orientations générales de la nouvelle Bibliothèque: accessible à tous, la Bibliothèque doit couvrir tous les champs de la connaissance, utiliser les technologies de l'information les plus modernes afin d'être consultée à distance, et entrer en relation avec les autres bibliothèques européennes. De surcroît, le Chef de l'État n'est pas intimidé par les polémiques: il les tranche lui-même ainsi qu'il fixe les délais de réalisation du projet, détermine les enveloppes budgétaires, choisit les personnes chargées de la conception et de la réalisation, etc. En définitive, la tradition de l'Ancien Régime semble respectée et réinvestie dans une Bibliothèque du troisième millénaire.

Le rayonnement culturel de la France dans le monde devient l'enjeu politique de la Bibliothèque. Comme aux lendemains de la Seconde Guerre mondiale, modernisation et indépendance nationale vont de pair⁵. Le Président annonce le projet avec des intentions "superlatives": "*Je veux que soit entrepris la construction et l'aménagement de l'une des ou de la plus grande et de la plus moderne bibliothèque du monde. [...] (i)l faut que Paris et que la France tout entière puissent démontrer ce qu'ils sont capables de faire*"⁶. Pour le Président, le caractère "moderne" de la Bibliothèque réside en grande partie dans les technologies de l'information. Il s'agit d'un défi à la modernité qu'il relève conjointement avec son *sherpa* de l'époque, Jacques Attali. Celui-ci a en tête la réalisation d'une bibliothèque immatérielle et consultable de tous points de l'Hexagone.

L'idée ne tombe pas du ciel. Les discours sur la "bibliothèque virtuelle" commencent à en faire rêver plus d'un. Les progrès d'une informatique de plus en plus performante, conviviale et abordable, de même que des innovations dans les réseaux de télécommunication, accompagnent des améliorations dans la numérisation de l'information elle-même, son stockage, son traitement et son accès. Dans cette évolution, le texte numérique est perçu comme l'impératif catégorique de la bibliothèque électronique où la numérisation de l'information (livres, manuscrits, images, sons...) prolonge l'accès

⁴ Cf. Simone BALAYÉ, "Un quadrilatère nommé Richelieu", in *La bibliothèque – Miroir de l'âme, mémoire du monde*, Paris, Autrement, avril 1991, pp. 81-92.

⁵ Cf. Robert GILPIN, *La science et l'État en France*, Paris, Gallimard, 1970.

⁶ François MITTERRAND, "Interview accordée par Monsieur François Mitterrand, Président de la République, à TF1 – Palais de l'Élysée, jeudi 14 juillet 1988", *Service de presse de la Présidence de la République*, p. 5. Une ambition qu'une part importante de la population semble juger légitime: 17% des Français considèrent le pays de la "Big Apple" comme une menace idéologique et culturelle. Cf. Alexandra SCHWARTZBROD, "Nucléaire: les Français sont pour la bombe mais contre les essais", *Libération*, 19-20 août 1995, p. 11

informatisé à l'information sur l'information (catalogues informatisés, banques de données documentaires...) ⁷.

Le projet présidentiel, se conformant à son désir de prestige, cherche à se situer en pionnier dans cette “*trajectoire technologique*”⁸: en 1988, on compte à peine 390 titres de CD-ROM chez le chef de file américain⁹ et 430 articles disponibles “en ligne” dans la base de données américaine DIALOG¹⁰. Toutefois, le discours des hommes d'État ne s'encombre pas de données techniques ou de statistiques. Pour justifier le projet technologique, démocratisation et diffusion du savoir sont des arguments politiquement plus envoûtants. Le rêve de Jacques Attali, c'est que, par ordinateur, Cahors puisse interroger Paris ou Washington en un clic de souris¹¹.

Dès lors, l'ambition de l'accès à distance s'institutionnalise. Il fait partie de la “charte” de la BNF et son Président n'imagine pas une bibliothèque du III^e millénaire sans les développements des technologies de l'information pour la conservation... et la consultation à distance des documents¹².

B. – DES PROPOSITIONS POUR LA TÉLÉTRANSMISSION DES DOCUMENTS

Pour légitimer un tel défi à la modernité et pour débloquer les crédits, outre l'invocation au progrès et les ambitions politiques et personnelles des dirigeants, on argue d'un “retard français”. L'argument, dans la bouche des professionnels du livre, a l'apparence de l'objectivité¹³. Patrice Cahart et Michel Melot¹⁴ ne manquent pas de l'évoquer dans leur rapport: au moment des *Propositions*, la France accuse un retard, disent-ils, dans le domaine des technologies de l'information, par rapport aux pays les plus avancés – États-Unis, Grande-Bretagne et Canada. Par exemple, contrairement aux États-Unis, la France n'a pas encore de catalogue collectif des monographies informatisé. De plus, la BN n'a pas complété l'informatisation de son catalogue tandis que celui de la

⁷ Cf. Jean-Pierre LARDY, *Les accès électroniques à l'information – État de l'offre*, Paris, ADBS Éditions, 1993, 90 p.; Alain GIFFARD, “Culture et éducation: l'invention d'un médium?”, in *La France à l'Exposition universelle (Séville, 1992)*, Paris, Flammarion, 1992, pp. 76-87.

⁸ La notion de trajectoire technologique est assimilée à la direction prévisible, à partir des exigences économiques, du développement des technologies. Cf. Christian DE BRESSON, *Comprendre le changement technique*, Ottawa, Presses de l'Université d'Ottawa, 1993, 386 p.

⁹ Veronica HARRY et Charles OPPENHEIM, “Evaluations of Electronic Databases, Part 1: Criteria for Testing CD-ROM Products”, *Online & CD-ROM Review*, 1993, vol. 17, n° 4, p. 211.

¹⁰ David EVERETT, “Full-Text Online Databases and Document Delivery in an Academic Library: Too Little, Too Late?”, *Online*, vol. 17, n° 2, mars 1993, p. 23.

¹¹ Cf. Thierry GANDILLOT, “La Bibliothèque du Président qui aimait les livres”, *Le Nouvel observateur*, 30 mars - 5 avril 1995, pp. 90-93.

¹² Cf. Jean FAVIER (entretien avec Bertrand Fisse), “Bibliothèque nationale de France, rencontre avec le Président”, *Officiel du livre*, pp. 34-36.

¹³ Le retard n'est-il pas plutôt une formule rhétorique que reprennent ceux qui veulent voir les bibliothèques se développer? L'argument du retard est d'ailleurs un thème récurrent dans l'histoire des bibliothèques françaises au XX^e siècle. Cf. Martine POULAIN, “Les bibliothèques françaises au XX^e siècle”, *Documentation et bibliothèques*, janvier-mars 1994, p. 5.

¹⁴ Patrice Cahart et Michel Melot sont les auteurs désignés par le Président des premières propositions sur la future Bibliothèque nationale de France. En 1988, Patrice Cahart est directeur des Monnaies et Médailles et Michel Melot est directeur de la Bibliothèque publique d'information du Centre Georges Pompidou.

Bibliothèque du Congrès (Washington) est terminé et celui de la British Library (Londres) est en voie d'achèvement¹⁵.

Néanmoins, le sentiment de retard est un des moteurs du changement: "*les retards n'empêchent ni les rattrapages ni surtout certains dépassements*"¹⁶. Les ambitions associées au projet de la BNF sont telles qu'elles suggèrent bel et bien le dépassement... et en particulier l'exploit technologique. L'usage des technologies de l'information et la notion d'accès à distance apparaissent comme la source et le levier d'une bibliothèque d'un "*type entièrement nouveau*"¹⁷. Le rattrapage serait en effet complet: informatisation entière du catalogue de la BN, création d'un catalogue collectif national informatisé à l'image des pays les plus avancés dans ce domaine, informatisation de la gestion de l'Établissement, création d'une collection numérisée et d'un outil de lecture assistée par ordinateur. Le programme de numérisation de la BNF, lui, est une des premières expériences de numérisation à échelle "industrielle" – 100 000 livres, soit près de 30 millions de pages. Le PLAO¹⁸, cette station de travail informatisée, spécialisée dans la lecture et la consultation des textes numérisés et que la BNF met à la disposition des chercheurs, constitue lui aussi une expérience nouvelle de consultation des textes électroniques: "*Aucune expérience de "bibliothèque électronique" n'avait jusque-là été menée suffisamment loin pour que soit validée ou infirmée l'idée de nouvelles lectures*"¹⁹.

Déjà dans le rapport Cahart-Melot, l'objectif novateur de la télétransmission, bien que toujours à l'honneur, ne semble pas suffire à rassasier les dirigeants. La transmission à distance cohabite dorénavant avec un nouvel objet, le PLAO, dont la justification ne se trouve pas du côté du marché, de la demande des utilisateurs ou de leurs besoins, mais bien du côté de l'offre ou de la volonté de réaliser une idée "*techniquement séduisante*"²⁰. "*Le PLAO répond au désir du Président de créer une "bibliothèque d'un type entièrement nouveau"*"²¹

¹⁵ Cf. Patrice CAHART et Michel MELOT, *Propositions pour une grande bibliothèque, Rapport au Premier ministre*, Paris, La documentation française, 1989, p. 81.

¹⁶ Jean-Jacques SALOMON, "*La capacité d'innovation*", in Maurice LÉVY-LEBOYER et Jean-Claude CASANOVA (sous dir.), *Entre l'État et le marché – L'économie française des années 1980 à nos jours*, Paris, Gallimard, 1991, pp. 15-58.

¹⁷ Mais les chiffres nuancent l'intérêt annoncé pour les technologies de l'information. Alors que cinq milliards de francs ont été consacrés à la réalisation du bâtiment – 69 % –, à peine 10 % ont été attribués aux technologies de l'information: 150 millions (2 %) sont accordés à la numérisation et au PLAO, 250 millions (3 %) vont à la réalisation du système d'information, et 150 millions (2 %) sont alloués au réseau documentaire de la BNF. La logique des Grands travaux commande en partie cette répartition. Comme pour le Grand Louvre ou l'Opéra Bastille, la priorité est donnée au contenant et à l'aspect matériel. C'est d'ailleurs un bâtiment vide que le Président de la République inaugure en mars 1995.

¹⁸ D'où vient l'idée de réaliser un poste de lecture? L'un des auteurs des *Propositions*, Michel Melot, a travaillé en 1988 dans le cadre de l'exposition "Mémoire du futur" avec le directeur de l'informatique de la BNF, Alain Giffard et le philosophe Bernard Stiegler, le "père" du PLAO.

¹⁹ Dominique CHOUCAN, "Du livre à l'ordinateur", *La Recherche*, n° 228, vol. 22, janvier 1991, p. 97.

²⁰ Le terme est emprunté à Christopher FREEMAN, "À quoi tiennent la réussite ou l'échec des innovations dans l'industrie", *Culture technique*, n° 18, mars 1988, P. 30-39.

²¹ AFP, "Un outil de travail nouveau pour les chercheurs conçu par la Bibliothèque de France: la "lecture assistée par ordinateur"", dépêche 221352, 22 mai 1990.

Encadré 1
Révolutionner la lecture
avec le poste de lecture assistée par ordinateur

Le caractère original du programme de numérisation dépasse la seule ambition de créer un outil "traditionnel" pour la consultation des documents électroniques. Dans l'optique de ses concepteurs, le PLAO doit servir un *nouveau* type de lecture et un *nouveau* contexte d'écriture. Les multiples possibilités offertes par l'outil sont mises en valeur dans une collection "enrichie", cohérente et encyclopédique. La numérisation n'est plus seulement conçue dans la perspective utilitariste de la télétransmission. Elle est aussi et surtout envisagée pour le développement d'une lecture assistée par ordinateur.

Développé pour une lecture savante, c'est-à-dire pour le public des chercheurs et non pas pour le grand public, le PLAO veut permettre un nouveau mode d'accès aux textes fondamentaux qui constituent la collection numérisée. Il ne s'agit plus seulement de faire du traitement de texte traditionnel (copier-coller, mise en pages, etc.) mais bien d'ouvrir la possibilité d'un travail de recherche sur le texte. Le PLAO permettrait la réalisation d'un nombre impressionnant de tâches²²:

- 1° d'obtenir en location à la BNF un ordinateur individuel, lié à la banque de données de la Bibliothèque et à des banques de données externes, qui permet une lecture et une écriture "assistées";
- 2° d'accéder aux catalogues, puis à la collection numérique, de la lire rapidement et de sélectionner les textes désirés;
- 3° de modifier le contenu de la collection numérisée en l'augmentant (pouvoir demander la numérisation de textes de la Bibliothèque, voire de textes extérieurs au fonds traditionnel);
- 4° de télécharger les ouvrages choisis et travailler indistinctement en mode "texte" ou en mode "image" (en passant de l'un à l'autre au besoin grâce à un logiciel de reconnaissance de caractères);
- 5° de lire les textes téléchargés à son rythme, pouvoir les annoter (oui! non!), retrouver facilement lesdites annotations (où ai-je relevé que Saint-Just avait dit la même chose?), surligner (rouge pour les jugements de valeur, vert pour les définitions, etc.), structurer (numéroter, séparer en parties, raturer), indexer (accrocher à certains mots une étiquette invisible qui permet ensuite de les trier, et d'en faire une table d'indexation), créer des liens hypertextes (un double-clic sur le mot Indépendance me renvoie automatiquement à Québec);
- 6° de retravailler les textes à sa guise, à l'aide de logiciels de traitement de texte évolués et de publication assistée par ordinateur (PAO ou *desktop publishing*); de les imprimer sur tout support (papier, transparents, diapositives) ou de les transférer;
- 7° d'organiser les textes en base de données (une fiche par chapitre, une autre fiche par personnage, une dernière par ouvrage), de manière graphique (organigrammes, arbres généalogiques, histogramme) ou textuelle (comparaisons inter- ou intra-textuelles, en lien avec les thésaurii, dictionnaires et autres ouvrages de référence);
- 8° de communiquer par courrier électronique avec l'administration de la BNF et les autres utilisateurs; de participer à des sessions de travail coopératif; d'accéder aux sites d'information (tableaux d'affichage électronique, fora).

Bref, il n'est quasiment rien que le PLAO ne puisse pas faire (sur papier tout du moins). Ainsi, pure spéculation, une chercheuse en histoire des sciences pourrait comparer le taux d'apparition du mot "force" dans les textes d'Archimède, de Bernoulli et d'Einstein, en déduire une loi de cooccurrence basée sur les χ^2 en fonction des contextes sémantiques, relever dans un fichier les passages les plus marquants pour défendre cette thèse, envoyer les données à Montréal à un collègue, imprimer le texte sur un transparent, et les graphiques sur une diapositive! Comment ne pas applaudir devant un tel programme?

²² Pour une description du PLAO, voir l'article de Jacques VIRBEL, "La lecture assistée par ordinateur et la station de lecture de la Bibliothèque de France", in A. VUILLEMIN, *Les banques de données littéraires*, Paris, Pulim, 1993, pp. 19-27.

Près de 70 PLAO seront mis à la disposition des chercheurs. Côté technique, ce sont des micro-ordinateurs dotés de disques durs de plusieurs centaines de méga-octets. Ils sont accompagnés de périphériques d'entrée et de sortie (*scanner*, lecteur de CD-Rom, imprimantes, écran haute définition, *modem*, etc.) et comportent des logiciels de lecture assistée et de reconnaissance de caractères. Enfin, ils sont intégrés au système d'information de la Bibliothèque, c'est-à-dire qu'ils peuvent communiquer avec les ordinateurs centraux où sont stockés en permanence les catalogues, les collections numérisées (textes, images, sons...), etc.

La création de PLAO y apparaît sous la forme d'un concept extérieur qui semble presque incident au départ. L'idée de postes de lecture et de recherche "*expérimentaux*" est évoquée en un paragraphe; les rapporteurs se limitent à une description fonctionnelle très générale. Néanmoins, les deux innovations sont annoncées:

Du point de vue de la communication, deux grandes nouveautés doivent marquer l'ouverture de la Grande Bibliothèque: la création, sur place, de postes de travail équipés d'instruments d'aide à la recherche, à la lecture et à la rédaction, et surtout la possibilité de consulter à distance des catalogues et des documents numérisés²⁵.

La télétransmission et le vœu initial d'une bibliothèque accessible à distance semblent donc toujours présider au destin de la BNF. Les rapporteurs ressassent les avantages du document numérisé: il peut être transmis de façon automatisée, il est accessible jusqu'au domicile des utilisateurs, il permet enfin des recherches de mots et d'occurrences à l'intérieur d'un document... L'accessibilité à distance est même glorifiée à l'occasion. On n'hésite pas à y voir mille possibilités nouvelles, et à faire des plans sur la comète technologique: les livres sont consultés à domicile sur l'écran de télévision de haute définition que chaque foyer, évidemment, possède²⁴.

Quelques incertitudes planent toutefois et elles ne sont pas passées sous silence: la nécessité d'effectuer des études de faisabilité technique et d'évaluation sociale est clairement notifiée. Entre autres surprises troublantes, il serait toujours possible que le lecteur refuse le livre numérique²⁵!

La méthode de numérisation est également soumise à la question. Dans l'optique de la télétransmission, Catherine Leloup présente une série de handicaps liés à une numérisation en mode "image" majoritaire²⁶. Les documents numérisés en mode "image" sont de dimension si considérable qu'ils sont chers à stocker et à transmettre et peu faciles à manipuler. La pondération mode "image" / mode "texte" trop favorable au premier s'avère nuisible à la mission de diffusion des savoirs en dehors des murs, bref, remettrait en cause, en partie au moins, les objectifs de télétransmission: *le risque essentiel est de constituer une base de documents primaires inaccessible ou difficile à utiliser, et par voie de conséquence, de bâtir un système image très coûteux car trop largement dimensionné²⁷.*

²³ Patrice CAHART et Michel MELOT, *op. cit.*, p. 99.

²⁴ Cf. Jean-Loup CHARRIER, "Une très grande bibliothèque numérique", in CAHART et MELOT, *op. cit.*, pp. 127-131.

²⁵ Il n'y aurait là rien de surprenant. L'histoire des techniques nous gave d'exemples d'apparent "refus du progrès": l'entreprise a refusé longtemps la béliographie avant de sombrer dans la faxo-mania; les spectateurs ont levé le nez, sans autre, pendant des années sur le vidéodisque des années 70 avant de tomber dans la nouvelle mode du CD-ROM multimédia.

²⁶ Cf. Catherine LELOUP, "Utilisation des techniques avancées de stockage et de transmission électronique pour une très grande bibliothèque", in CAHART et MELOT, *op. cit.*, pp. 133-155.

²⁷ *Ibid.*, p. 155.

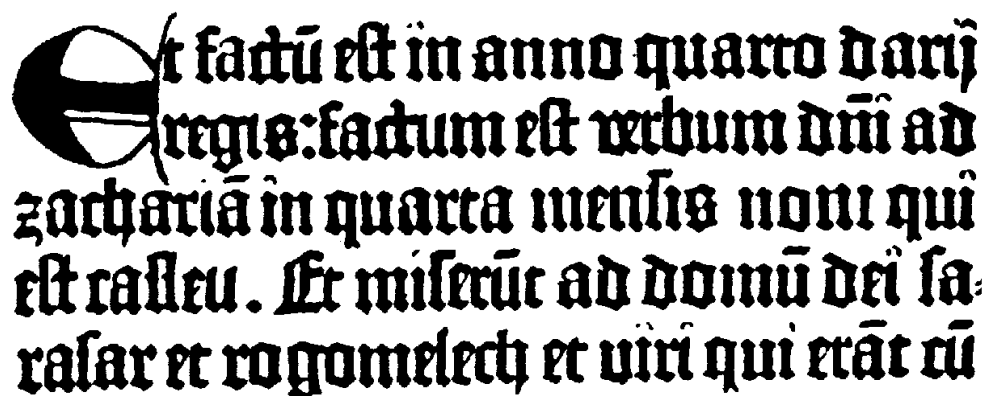
Encadré 2

Le dilemme mode "image"/ mode "texte"

Le processus de numérisation n'est pas exempt de choix reposant tantôt sur des caractéristiques techniques, tantôt sur des considérations juridiques, pratiques, économiques et intellectuelles. Le choix d'une numérisation en mode "image" ou en mode "texte" est lui-même déterminant pour l'usage de la collection numérisée.

La numérisation des textes se fait grâce au numériseur, couplé à un ordinateur: une barrette de petits yeux électroniques (des diodes photoélectriques) scrute ligne par ligne les pages d'un livre massicoté ou d'un microfilm; les caractères sont transformés en une série de points, puis convertis en codes numériques dans un fichier informatique. En fait, la machine enregistre la présence ou l'absence d'encre et attribue à chaque point la valeur 0 ou 1 selon qu'il y a ou non de l'encre. Ce procédé est celui de la numérisation en mode "image" – qu'on appelle aussi *bit-map*. Ce qui est numérisé en mode "image", c'est l'image de la page. Le résultat est une représentation fidèle de la page imprimée: sa mise en pages mais aussi sa typographie, et même les défauts du papier ou du parchemin...

Un extrait de la *Bible* à 42 lignes numérisé en mode "image"



Et factū est in anno quarto darij
regis: factum est verbum dñi ad
zachariā in quarta mensis noni qui
est casleu. Et miserūc ad domū dei sa
rasar et rogomelech et viri qui erāt cū

Source: GENFLEISCH, Johann (dit GUTENBERG), 1452. *Bible mazarine, ou Bible à 42 lignes*, Mayence: Gutenberg. Illustration tirée de la Collection Gaston-Burnand, in FABRE, Maurice, 1964. *Histoire de la communication*, Lausanne (Suisse): Éditions Rencontre & Erik Nitsche International, p. 41 (105 pages).

Inconvénient de cette grande fidélité à l'original, les caractères ne sont pas "reconnus" automatiquement par l'ordinateur. Pour effectuer des analyses sur un texte – des calculs d'occurrences, des recherches de mots, du traitement de texte, par exemple – il faut non seulement que le livre soit numérisé, mais en plus que l'ordinateur soit muni d'un logiciel de reconnaissance optique de caractères (ROC, ou OCR en anglais). Grâce à ce logiciel, on accomplit la seconde étape qui consiste à convertir le document du mode "image" au mode "texte". Longue et fastidieuse en raison du nombre important d'erreurs de reconnaissance (il suffit d'un changement de fonte ou d'un défaut d'encre pour que l'ordinateur soit... dépassé par les événements), cette étape consiste pour l'ordinateur à retravailler avec l'image de la page précédemment numérisée: il la découpe (virtuellement) en zones de texte (négligeant dessins et imperfections du support), puis il compare chaque élément typographique aux caractères d'un alphabet qu'il a en mémoire, et conclut en transcrivant enfin le caractère typographique qu'il croit avoir deviné dans un nouveau document informatique. Au bout de cette chaîne, en théorie tout du moins, l'enluminé *Bonjour* devient un chiche *Bonjour*. Pour les textes "anciens", la reconnaissance des caractères offre cependant des performances encore limitées. Dans un monde parfait, nous obtiendrions le résultat suivant si nous numérisions en mode "texte" l'extrait précédent.

L'extrait précédent de la *Bible* à 42 lignes idéalement numérisé en mode "texte"

Et factum est in anno quarto Darii regis: factum est verbum domini ad Zachariam in quarta mensis noni qui est Casleu. Et miserunt ad domum dei Sarasar et Rogomelech et viri qui erant cum

La réalité dévoile cependant une technique encore bien imparfaite. Un document sans typographie complexe et numérisé en mode "texte" présente des défauts agaçants.

Exemple d'une numérisation réelle en mode "texte"

10

LE GESTE ET LA PAROLE

leurs rapports. Cette opinion transparait dans les traditions populaires européennes, dans des contes où les animaux parlent et s'insinuent dans le déroulement des comportements humains.



10 LE CEST~. ET I~ P~ROLE

leurA rapports. Cette opinion transparait dans les traditions populaires européennes, dans des contes ol`l leA animaux prlent et s'insinuenl dans le déroulemellt des com rte ments humains.

Source: AITIC, *État de l'art sur la lecture optique de pages*, document interne, octobre 1990, non paginé.

Le mode "texte" pose d'autres problèmes; légal, par exemple. La mise en forme originale n'est plus respectée, on l'a vu. Or ce dernier aspect est essentiel pour respecter l'obligation légale de la reproduction fidèle de la mise en page originale du livre qui n'est pas encore tombé dans le domaine public. Ou économique: le mode "texte" est désavantagé par rapport au mode "image". Ce dernier évitant l'étape de la reconnaissance des caractères, son coût peut être dix à cinquante fois moins cher qu'une numérisation en mode "texte".

En revanche, la numérisation en mode "image" est très coûteuse en terme de stockage ou de transmission. Un fichier en mode "image", même compressé par des logiciels spécialisés, occupe beaucoup plus d'espace en mémoire qu'un fichier en mode "texte". Par exemple, un livre de 300 pages numérisé à la BNF et compressé occupe 20 millions d'octets en mode "image", ce qui est considérable par rapport à l'équipement informatique standard d'un usager. Numérisé en mode "texte" et dans certaines conditions, le même livre occuperait un espace d'à peine 600 mille octets. En conséquence, le coût de stockage des fichiers en mode "image" est plus élevé, le temps de transmission est plus long, la communication est moins fiable et les ordinateurs manipulant ces textes doivent être performants. Des frustrations que les usagers d'Internet connaissent quand ils téléchargent des images sur leur ordinateur. Des réseaux de télécommunication à haut débit étendus à l'échelle nationale permettront cependant de faciliter la transmission de fichiers image lourds.

Autre avantage inhérent au mode "texte": la structuration des documents en fonction d'une utilisation efficace appelée "hypertexte": en particulier, dans les dictionnaires, les encyclopédies, etc., un double-clic avec la souris sur un mot à l'intérieur d'une définition renvoie automatiquement à la définition dudit mot... Un mot suffit à mettre en relation plusieurs textes à la manière des pages WEB. Un usage que la numérisation en mode "image" n'autorise pas automatiquement.

Dès les premières études, la télétransmission est donc un enjeu central de la Bibliothèque. Officiellement elle le reste, mais les choix qui vont être faits pour le programme de numérisation vont peu à peu donner une place relative à cet objectif.

C. – UN DÉTOURNEMENT PAR LA BANDE

L'objectif n'est pas remis en cause ni attaqué "de front"; mais il va devenir peu à peu la variable dépendante du programme de numérisation, dont il n'est bientôt presque un aboutissement parmi tant d'autres. On peut donc parler d'un détournement par la bande. Des considérations à la fois techniques, économiques et intellectuelles ont mené la Bibliothèque à faire s'éloigner, petit à petit, le programme de numérisation de l'objectif de télétransmission, et à implanter des PLAO en lieu et place de celle-ci.

Les réflexions des concepteurs ont débouché, en effet, sur un ensemble de principes intellectuels auxquels ils entendent se tenir fermement. La collection numérisée est constituée *a priori* selon les critères de cohérence et d'encyclopédisme.

À la différence du programme CARL²⁸ aux États-Unis ou de l'expérience EDIL²⁹ en Europe où les textes sont numérisés à la demande, la BNF numérise les documents *a priori*. Ce sont des équipes spécialisées dans diverses disciplines qui décident du contenu de la collection associée à leur propre domaine de recherche.

Encadré 3

Exemples de programmes de numérisation concurrents

CARL, une expérience américaine

Le cas du programme CARL³⁰ permet de comparer l'expérience française de la BNF à un modèle de changement américain. Celui-ci est décentralisé, fondé sur la demande et sur des adaptations progressives; celle-là est centralisée, fondée sur l'offre et sur des changements qui surviennent par à-coup.

À l'origine, le programme CARL est l'initiative d'un directeur de bibliothèque qui pensait que si les petites bibliothèques de recherche du Colorado s'associaient et combinaient leurs ressources, elles pourraient ensemble fournir la documentation d'une grande bibliothèque de recherche. Le programme débute en 1974. Il s'agit alors d'un groupe informel de directeurs de bibliothèques du Colorado intéressés par la coopération dans le but de "*favoriser le progrès et l'avancement des bibliothèques de recherche*" de l'État. En 1978, le groupe informel devient une petite association à but non lucratif. En 1981, le système informatique CARL vient fédérer les collections des bibliothèques membres en permettant de les gérer comme une seule entité. Il permet la consultation des catalogues et la réalisation des tâches administratives, telle la circulation des documents, leur acquisition, le contrôle des périodiques et un suivi bibliographique.

En 1981, CARL comporte 77 terminaux. Dix ans plus tard, 1 000 terminaux sont directement liés au serveur de Denver. En plus de relier les bibliothèques entre elles, le réseau est également connecté à tous les campus universitaires

²⁸ Colorado Alliance of Research Libraries

²⁹ Electronic Document Interchange between Libraries

³⁰ Pour une petite histoire du programme CARL, voir Rebecca T. LENZIZI et Ward SHAW, "Creating a New Definition of Library Cooperation: Past, Present and Future Models", *Library, Administration & Management*, vol. 5, n° 1, hiver 1991, pp. 37-40.

du Colorado, et à Internet. Plus de 50 bases de données forment le système, qui donne ainsi accès à plus de 30 catalogues de bibliothèques et à un grand nombre de bases de données, dont l'encyclopédie Grolier en texte intégral.

Au système CARL s'ajoute le service d'accès aux périodiques UNCOVER. En octobre 1990, cette base de données comporte plus de 10 000 titres de périodiques. L'enregistrement du titre du périodique dans UNCOVER est immédiat. Le catalogue ainsi constitué, première étape du développement de UNCOVER, permet à toute personne de savoir si une bibliothèque du réseau possède un certain périodique. Cela évite aux bibliothèques du réseau d'acquérir les mêmes journaux.

UNCOVER prévoit aussi le développement des outils permettant la transmission des documents eux-mêmes aux usagers. Les usagers consultent les résumés des articles, les commandent sur le champ auprès de l'éditeur ou d'une tierce organisation et se les font envoyer moyennant finance pour couvrir les frais de photocopie et les *copyrights*. Ils peuvent également – troisième étape du développement d'UNCOVER – demander l'article "en ligne".

Pendant l'année 1990, le système CARL a commencé à fournir des textes intégraux d'articles de périodiques. Le lecteur communique alors les références de l'article désiré, son code de carte de crédit et un numéro de fax. Quelques minutes plus tard, voici l'article tout chaud sorti des entrailles de la machine.

Éventuellement, le système UNCOVER projette d'offrir un service de numérisation à la demande. Il suffira alors à l'utilisateur de commander l'article et de le prépayer, puis de le retrouver sous sa forme numérique à la bibliothèque la plus proche. L'article numérisé peut ensuite être transmis à un centre d'entreposage optique où il est stocké pour les futures demandes.

EDIL, une expérience européenne

Plus jeune et plus courte que l'expérience CARL, l'expérience européenne EDIL³¹ commence en 1993 pour se terminer en 1995. Elle envisage aussi la numérisation des documents comme un moyen de transmission des documents à distance. Financé par la Communauté européenne et par les pays participants (Allemagne, Grande-Bretagne, France et Hollande), EDIL a comme objectif de vérifier s'il est possible d'interconnecter des bibliothèques ayant recours à des infrastructures de télécommunication différentes. Les bibliothèques participantes ont pour tâche d'échanger des documents par voie numérique.

La bibliothèque emprunteuse identifie le document demandé par le lecteur. Elle dépose une demande de prêt par messagerie à la bibliothèque prêteuse. Celle-ci n'a pas le document en sa possession: elle en avise la bibliothèque emprunteuse par messagerie. Au contraire, si la bibliothèque prêteuse détient le document, elle le numérise et l'envoie à la bibliothèque emprunteuse par transfert informatique de fichiers. Aussitôt transmis, le document numérisé est détruit. Des expériences antérieures auraient en effet relevé que le pourcentage des documents transmis et redemandés est extrêmement faible.

Deux grandes lignes directrices guident leur choix. Le premier est celui d'une cohérence de type éditorial. Aux yeux des concepteurs, les bases de données actuelles, fondées sur le seul critère de la demande, sont conçues comme des "sacs de billes"³², accumulant les ouvrages les uns "sur" les autres. Du coup, l'utilisateur ne sait jamais si le document qu'il désire fait partie de la collection numérisée ou non. On peut remédier à cette lacune, et favoriser la visibilité et l'utilisation de la collection, en établissant un principe intellectuel de cohérence. Elle devrait permettre une "appropriation appro-

³¹ Claude JOLLY, "Les projets ION et EDIL", in *Transmission électronique du document*, op. cit., pp. 37-49.

³² Alain GIFFARD, entretien, 7 mars 1995.

*fondie*³³ de la collection numérisée, que ce soit sur place ou à distance. La cohérence est également définie en fonction du contenu de la collection de la BN. La collection numérisée regroupera des ouvrages de base et textes de référence... à l'image de la BN: c'est une collection de classiques numérisés dans une bibliothèque "*classique*"³⁴. De fait, la collection numérisée ressemble à un corpus de références constituant le fonds commun de travail des chercheurs et le cœur du patrimoine culturel que conserve la Bibliothèque. La collection regroupe des œuvres classiques et des ouvrages de références – dictionnaires, almanachs, encyclopédies, etc. – en littérature, philosophie, histoire, sciences humaines, science politique, science juridique, science économique et histoire des sciences. Quelques périodiques enfin viennent aérer la collection, composée essentiellement de monographies³⁵.

Encadré 4 **Les collections de la BNF**

On retrouvera quai de Tolbiac:

- 10 000 000 d'imprimés en fonds patrimonial et 400 000 volumes en libre accès;
- 350 000 titres de périodiques en fonds patrimonial;
- 123 000 microfilms et 1 400 000 microfiches;
- 100 000 textes numérisés;
- 300 000 images fixes numérisées;
- 440 000 heures de documents sonores;
- 70 000 heures d'images animées.

Resteront rue de Richelieu:

- 880 000 cartes et plans;
- 11 000 000 d'estampes et de photographies;
- 530 000 manuscrits;
- 580 000 monnaies, médailles et antiques;
- 2 100 000 pièces et recueils de musique;
- 3 000 000 d'affiches, photos, manuscrits, maquettes, recueils de presse dans les arts du spectacle.

Les collections d'imprimés seront organisées en quatre départements thématiques: 1) Philosophie, histoire, sciences de l'homme et de la société; 2) Sciences politiques, économiques et juridiques; 3) Sciences et techniques; 4) Littérature et art. En fait, la BNF est avant tout une bibliothèque de sciences humaines. Bien qu'elle bénéficie du dépôt légal en France, elle limite ses acquisitions étrangères aux humanités et plus précisément à l'histoire et à la littérature. Côté lecteurs, la répartition des domaines de recherche semble

³³ Yannick MAIGNIEN, "La constitution de la collection numérisée de la Bibliothèque de France: vers un nouvel encyclopédisme?", communication au colloque *La nouvelle technologie au service de la littérature*, Oxford, 1-3 octobre 1993, 14 p.

³⁴ *Ibid.*

³⁵ Pour exemple, le corpus d'histoire des sciences regroupe (en 1993) environ 400 titres, choisis dans la vaste période entre l'Antiquité et... le XX^e siècle. Il comprend entre autres des œuvres d'Archimède, Euclide, Hippocrate, Ptolémée, Al-Fakhri, Brahé, Copernic, Vinci, Bernoulli, Galilée, Leibniz, Lavoisier, Ampère, Arago, Lamarck, Volta, de Broglie, Einstein, Koyré, Peirce, Borel, Heisenberg, etc. Quelques biographies scientifiques émaillent également la collection, ainsi que des études récentes sur une partie des œuvres ou des essais historiques et philosophiques sur la discipline. Et pour l'actualité des débats, on trouve quelques *Comptes rendus et mémoires de l'Académie des sciences*, *Philosophical Transactions of the Royal Society*, etc. (Michel RICHARD, "Le programme de numérisation de la Bibliothèque de France", *Bulletin des Bibliothèques de France*, t. 38, n° 3, 1993, p. 53-63.) Mais l'œcuménisme scientifique doit attendre; le fonds est d'abord spécialisé en sciences pures!

correspondre à l'orientation générale du fonds documentaire: 37% effectuent des recherches en histoire, 22% en littérature et 20% en histoire de l'art, 5% en philosophie, 3% en sociologie, ethnologie, 2% en linguistique et 6% travaillent sur d'autres sujets³⁶.

Le deuxième principe est un encyclopédisme "immédiat". Dès l'ouverture, le fonds numérisé doit être assez exhaustif pour répondre aux besoins en recherche dans chacune des disciplines de la BNF. Cela impose une contrainte quantitative: en 1990, on parlait d'un million de documents numérisés. L'ambition est vite jugée irréaliste compte tenu du coût et du temps requis. On fixe alors un "seuil encyclopédique" à 300 000 documents, deuxième objectif dont viendra à bout la seconde "cohabitation" politique pendant laquelle on réduira le nombre à 100 000. Cela impose également une contrainte qualitative: la collection ne se veut plus simplement fondée sur l'"accumulation" des documents, elle cherche à être "*un ensemble d'outils scientifiques intellectuellement sophistiqués*"³⁷. D'où la volonté de couvrir tout de suite toutes les périodes historiques, et tous les domaines.

Quelles sont les implications de ces choix intellectuels sur les choix techniques? On peut immédiatement établir une relation entre les deux principes, cohérence et encyclopédisme, et le choix de la numérisation en mode "image". Cela va de pair! D'une part, la numérisation en mode "image" en offrant un coût plus faible, permet d'appliquer le procédé à un plus grand nombre de textes qu'en mode "texte" – avantage incomparable au titre de l'encyclopédisme et de la cohérence, et désavantage important pour la future télétransmission. Pour l'utilisateur à distance qui possède un micro-ordinateur "standard", le document numérisé en mode "image" est plus long à transmettre, plus lourd à entreposer et plus difficile à manipuler que son équivalent en mode "texte". D'autre part, la numérisation en mode "texte" des documents anciens offre des performances moins satisfaisantes – ce qui, *a contrario*, favorise là encore, le mode "image", apparemment plus conforme à l'idée d'une collection numérisée encyclopédique. Tout concourt à faire une collection numérisée qui puisse difficilement, dans l'état actuel de la technique, être télé-accessible.

Compte tenu aussi du temps limité dont dispose la BNF pour numériser sa collection (la Bibliothèque aurait dû ouvrir avant la fin du deuxième septennat de François Mitterrand), le mode "image" est privilégié. C'est donc 90% de la collection qui seront accessibles directement en mode "image" et seulement 10% (les ouvrages de références pour lesquels la recherche par mots ou occurrences est essentielle) qui seront accessibles directement en mode "texte".

Quid de la télétransmission des documents? Peu à peu, cet enjeu semble avoir du plomb dans l'aile. La consultation des documents numérisés apparaît de plus en plus tributaire d'un outil de consultation ultra-performant que veut être le PLAO. Car dans un "système technique"³⁸ où "*l'extension de la digitalisation de l'information va [...] de pair avec le développement de la filière électronique, le recours croissant aux réseaux de télécommunications et la montée du secteur des logiciels*"³⁹, l'accès et la transmission à distance des documents numérisés en mode "image" souffrent des incohérences

³⁶ Christian BAUDELLOT et Claire VERRY, "Profession: lecteur?", *Bulletin des bibliothèques de France*, t. 39, n° 4, 1994, pp. 8-17.

³⁷ Alain Giffard, cité par Yannick MAIGNIEN, *op. cit.*, p. 2.

³⁸ Le concept de Bertrand Gille se fonde sur les idées de l'interdépendance des techniques et sur la cohérence nécessaire de ces techniques entre elles. Cf. *Histoire des techniques*, Paris, Gallimard, "Bibliothèque de la Pléiade", 1978.

³⁹ Philippe BRETON, *Une histoire de l'informatique*, Paris, La découverte, 1990, p. 216.

techniques. Malgré les moyens de compression de fichiers, l'espace qu'occupent les documents en mode "image" demeure considérable. Ainsi, les conditions d'accès sont difficiles pour le système informatique du lecteur à domicile qui doit avoir à sa disposition un écran d'une bonne définition, une capacité de stockage suffisante pour entreposer un ou plusieurs documents de 20 mégaoctets, une capacité de traitement assez élevée pour manipuler le texte à sa disposition, des logiciels appropriés pour exécuter différentes fonctions sur le texte, une imprimante de qualité afin de rendre sur papier une résolution suffisante pour une lecture agréable, etc.⁴⁰ De plus, les problèmes que pose la communication des documents en mode "image" persistent. Le directeur de la DINT⁴¹ estime que "les débits des réseaux disponibles actuellement ne sont pas suffisants". Idéalement, il faudrait "disposer de liaisons à haut débit pour éviter qu'il faille une demi-heure et plus pour transmettre un bouquin entre Paris et Marseille"⁴². Encore aujourd'hui, la transmission s'avère relativement longue et coûteuse pour le lecteur de la Cannebière!

Dans ces conditions, l'échéance d'un accès à distance s'éloigne toujours davantage. Le constat est fait dès 1990⁴³: les textes numérisés ne pourront être accessibles autrement qu'entre les murs de la Bibliothèque, grâce aux postes de lecture. Une étude prévoit avec prudence la télétransmission en mode "image" pour 1994!⁴⁴

Le pari qu'a fait la BNF en misant sur l'évolution et la diffusion des technologies de l'information retarde l'échéance de l'accès à distance au texte. Dans les années à venir, l'évolution des technologies lui donnera-t-elle raison? C'est possible. À condition que le contenu de la collection numérisée réponde aux besoins des usagers à distance. Des désaccords existent à ce sujet. Par exemple, la télétransmission est-elle appropriée pour les textes longs comme pour les textes courts⁴⁵? Les textes numérisés, s'ils sont disponibles dans d'autres bibliothèques en France, seront-ils consultés électroniquement⁴⁶? Encore faudra-t-il que la lecture assistée par ordinateur, devenue petit à petit l'élément central de la collection numérisée, soit adoptée par le lecteur!

D. – UN NOUVEAU MÉDIUM POUR LA LECTURE

Ainsi le futuriste PLAO se destine-t-il à être essentiel à la consultation de la collection numérisée. Il se veut un outil suffisamment performant pour traiter les textes

⁴⁰ Étant donné les problèmes de transmission et de compression des images, les concepteurs du projet *Beowulf* de la British Library ont finalement opté pour l'option CD-ROM plutôt que la télétransmission.

⁴¹ Direction de l'informatique et des nouvelles technologies

⁴² Cf. André ZYSBERG cité par Jean-François AUGEREAU, "La lecture à l'heure de l'informatique", *Le Monde*, 28 mars 1995, p. 21.

⁴³ Cf. Jean GATTÉGNO, *op. cit.*

⁴⁴ Cf. EUTELIS, *Étude de faisabilité d'un service pilote de télé-consultation et de télé-communication d'ouvrages numérisés de la Bibliothèque de France*, 18 décembre 1991.

⁴⁵ Pour Hervé LE CROSNIER, "ce qui est en jeu dans le réseau, ce n'est pas la transmission de textes longs [plus facile à acheter en format imprimé], ce n'est pas la transmission de romans, ce n'est pas la transmission de ces textes qui font notre culture, qui font notre base de références, mais la transmission de textes courts qui sont des articles de recherche scientifique". Cf. *Transmission électronique du document*, Paris, Bibliothèque publique d'information, 1994, p. 144.

⁴⁶ Dans le cadre du rapport Miquel sur la BNF déposé en 1992, Roger Chartier insiste sur le critère de rareté des documents numérisés. Peu importe que les textes soient longs ou courts, ce qui compte c'est qu'on ne les trouve pas ailleurs qu'à la BN. Cf. COMMISSION "BIBLIOTHÈQUE DE FRANCE", *Recueil des attendus adoptés*, novembre 1992.

numérisés et il est inextricablement lié à la collection numérisée et vice-versa: “*les postes de lecture n’ont pas de sens sans données textuelles informatiques et les fonds numériques sont sous-utilisés sans terminaux d’accès*”⁴⁷.

Ce lien congénital qui unit le contenant et le contenu ne fait cependant pas d’emblée l’unanimité. Pour le Président du conseil scientifique de l’EPBF⁴⁸, l’historien du livre Roger Chartier, cette réflexion parie sur une évolution incertaine des pratiques de lecture. L’historien défend l’idée qu’il eut été plus conforme aux évolutions majeures dans l’histoire de l’imprimé de mettre l’accent sur la transmission à distance et de concevoir le programme de numérisation en fonction de la demande de communication des documents ou de leur fragilité⁴⁹. Mais, en dépit de cette réserve, isolée dans l’océan d’enthousiasme qui baigne le PLAO, et plus sensible à l’évolution des pratiques et des usages actuels des nouvelles technologies de l’information qu’à l’évolution possible de l’offre technique, l’objectif demeure de développer une collection numérisée associée à un outil de consultation et de lecture.

D’un autre côté, le soudain intérêt pour ces drôles de machines, qui n’avaient pas encore été inventées⁵⁰, offre un certain nombre d’avantages incidents, et en premier lieu la création d’un nouveau type de lecture.

*[C]e n’est pas seulement le livre comme objet parcouru qui serait le support de la lecture – naturelle ou assistée –, mais bien un être plus complexe, plus diffus, plus délicat à définir en tant qu’agrégats de documents*⁵¹.

Bien que cette conception “moderniste” de la lecture électronique ne soit pas partagée par tous au sein des différentes “communautés intellectuelles”, les membres de l’institution s’y associent les uns après les autres. L’évolution des technologies de l’information leur semble suffisante “*pour que deviennent imaginables de nouvelles pratiques d’écriture et de lecture*”⁵². Dès lors, ce n’est plus l’accessibilité à distance qui importe; ce sont surtout les nouveaux outils de consultation.

La Bibliothèque de France est avant tout une bibliothèque de recherche très moderne de dimension internationale. À l’évidence, sa modernité tient tout d’abord à l’utilisation qu’elle fait des technologies numériques d’information. Plus particulièrement, cette bibliothèque conduit à envisager non seulement un nouveau type de consultation des fonds (et donc une nouvelle bibliothéconomie), mais aussi un nouveau type de lecture et du même coup un nouveau contexte pour l’écriture, s’il est vrai qu’une bibliothèque de recherche est un lieu de travail.

C’est en ce sens que le rapport de Patrice Cahart et Michel Melot décrit une consultation des catalogues, mais aussi, dans certaines conditions, des

⁴⁷ Cf. Jean GATTÉGNO, *op. cit.*

⁴⁸ Établissement public de la Bibliothèque de France

⁴⁹ Cf. Jean GATTÉGNO, *op. cit.*

⁵⁰ Les machines à lire sont conçues par un groupe de travail sur le PLAO formé en 1989. Le groupe est présidé par un philosophe, Bernard Stiegler, maître de conférence à l’Université technologique de Compiègne. Il réunit aussi des chercheurs au CNRS dans le domaine de l’informatique et de la linguistique, et des spécialistes en documentation. L’équipe compte sept personnes.

⁵¹ Yannick MAIGNIEN, “La lecture assistée par ordinateur. Lector ex machina”, *Le Débat*, n° 86, septembre-octobre 1995, p. 155.

⁵² Bernard STIEGLER, “Machine à lire”, in *La Bibliothèque*, Paris, Autrement, n° 121, avril 1991, p. 143.

*fonds eux-mêmes, sur consoles électroniques, et surtout, une consultation étendue jusqu'au traitement des documents ainsi rendus accessibles*⁵³.

Comme pour la collection numérisée, les réalités techniques ne commandent pas tout dans la conception du PLAO. Chez ses concepteurs, un biais "intellectuel" milite pour un "rapprochement plus fondamental de l'informatique et des sciences et savoirs du texte"⁵⁴. Les fonctionnalités imaginées pour le PLAO reposent sur les connaissances du groupe de travail en matière de lecture savante: on parle vaguement d'une lecture d'"investigation", qui se réalise dans la durée, qui combine lecture et écriture, et qui repose sur des hypothèses de recherche. On glose aussi sur un travail de "libre créativité". Sur un plan technique, on cherche à faire coïncider les mêmes fonctions de lecture – sauf bien sûr pour ce qui concerne la reconnaissance des caractères – à un document numérisé en mode "image" ou en mode "texte".

Pour valider l'outil, des prototypes sont réalisés et testés par... sept chercheurs solitaires. C'est en interrogeant ces chercheurs sur leur nouvelle expérience que le choix du mode "image" est confirmé. Les observations des cobayes ont été, en effet, très favorables aux stations manipulant du mode "image"⁵⁵. Pour une raison: "La lecture en elle-même, dans un très grand nombre de cas, est plus facile en mode image, fidèle à la typographie, que dans le mode texte"⁵⁶.

La réussite du PLAO demeure en grande partie tributaire d'une évolution incertaine de l'offre technique. Quelques exemples⁵⁷: la capacité des mémoires et des unités de stockage, le développement de logiciels documentaires "permettant à l'utilisateur de structurer lui-même ses propres bases de données [...] de plus en plus souvent en texte intégral", l'intégration d'éléments nouveaux dans les bases documentaires comme l'annotation, la création de liens hypertextuels, etc., et l'affinement du savoir logiciel de l'informatique documentaire en matière de traitement statistique du texte intégral... De plus, l'offre technique doit être adaptée aux particularités du PLAO. Il semble que les outils développés le soient principalement pour les professionnels de la documentation. Or leurs usages sont différents de ceux des lecteurs visés par le PLAO: qui se servira vraiment de toutes les caractéristiques offertes par ces petits bijoux de sophistication, hors du temps, hors des normes habituelles? Les spécialistes en informatique ou les chercheurs qui auront la patience de prendre plusieurs dizaines d'heures de formation pour utiliser cet outil singulier? À moins que les usagers aient bientôt à leur disposition des outils commercialisés du même ordre que ceux proposés sur le PLAO. Les évaluations des concepteurs sont positives: l'appropriation de la micro-informatique par les usagers évolue, le multimédia se diffuse dans les foyers, la numérisation se répand de plus en plus, la capacité des micro-ordinateurs augmente... suffisamment pour conclure au développement du PLAO, qui ne devrait plus bientôt être la seule marque de la BNF. Mais espérons qu'à ce moment là, le PLAO ne sera pas dépassé par les technologies commercialisées⁵⁸.

⁵³ BIBLIOTHÈQUE DE FRANCE, *Les postes de lecture assistée par ordinateur de la Bibliothèque de France, rapport d'activité et recommandations du groupe de travail*, janvier, 1990, p. 4.

⁵⁴ Alain GIFFARD, *op. cit.*, p. 87.

⁵⁵ Cf. Fabien GRUIER, "Quand les puces auront mangé les livres", *Le Nouvel observateur*, 30 septembre - 6 octobre, 1993.

⁵⁶ Alain GIFFARD et Michel RICHARD, *Programme de numérisation, le projet et la collection*, 1^{er} septembre 1992, p. 28.

⁵⁷ Cf. BIBLIOTHÈQUE DE FRANCE, *Les postes de lecture assistée par ordinateur de la Bibliothèque de France, op. cit.*

⁵⁸ Le développement du visiophone aboutit au même genre de problème: une fois les difficultés techniques résolues, c'est l'intérêt même de la technologie qui est remis en cause dans un environnement où les micro-ordinateurs sont interconnectés. Cf. Frédéric BARBIER et Catherine

E. – DE LA MENACE POUR LES ÉDITEURS À LA MENACE DES ÉDITEURS

Il n’y a pas que la technique qui résiste aux idées volontaristes. Aux problèmes de “quincaillerie” s’ajoutent des problèmes légaux. Car les concepteurs du programme de numérisation, en voulant bien faire sous prétexte de cohérence et d’encyclopédisme, ont inclus dans leur collection numérisée idéale des ouvrages dont les droits ne sont pas encore “tombés”.

*L’ambition d’offrir à l’utilisateur l’ensemble d’un corpus concernant un auteur, toutes les traductions de référence, toute la critique qui fait autorité, la ou les meilleures éditions d’œuvres complètes, etc. conduit à prendre ou sélectionner dans les livres mis actuellement sur le marché tous ceux qui répondent à ce critère*⁵⁹.

Or la France est signataire de la Convention de Berne sur le droit d’auteur qui prévoit que la “*propriété littéraire dure donc, en règle générale, pendant la vie de l’auteur et cinquante après sa mort*”⁶⁰. Le temps de protection a été récemment harmonisé par la Communauté européenne, et porté à 70 ans après la mort des “gens de lettres” (non comprises certaines prolongations). Donner en libre accès le contenu de livres couverts par le droit d’auteur, c’est enlever le pain de la bouche des auteurs et éditeurs! La BNF en est consciente. Elle y dédie même une rubrique sur son site Internet et rappelle aux usagers les règles de la propriété littéraire.

Des 100 000 documents numérisés par la BNF, un peu plus de 30 000 (représentant un investissement d’environ 30 millions de francs) sont concernés par le droit de l’éditeur et de l’auteur! La BNF doit donc négocier avec le monde éditorial de l’accessibilité à distance, du téléchargement sur disquette ou autre support numérique et même de la simple consultation d’un texte sur le PLAO. Si des ententes ne sont pas conclues, la partie concernée de la collection soumise au droit de consultation et de reproduction risque purement et simplement d’être amputée. C’est le symptôme du détournement de la *mission patrimoniale* vers une *mission éditoriale*: aussitôt, les éditeurs font barrage.

Pour ces derniers, l’enjeu n’est pas nouveau. Ils s’estiment déjà mis en péril par la mission de communication des bibliothèques publiques, et revendiquent une meilleure rétribution et un meilleur contrôle de la photocopie et l’instauration du prêt payant des livres hors domaine public⁶¹. “*Copier est devenu trop facile*”⁶², clame le Syndicat national de l’édition. Et le libre accès à une collection numérisée n’est pas pour les rassurer. Pour eux, tous ces phénomènes contribuent au non-achat de l’ouvrage. Bref, *a priori*, rien ne prédispose les éditeurs à devenir les partenaires de la BNF. D’autant que les effets de la numérisation du livre sont d’une autre échelle: un livre numérisé peut être reproduit indéfiniment, dans son ensemble ou en partie, dans le monde entier *via* les réseaux informatiques! C’est sans commune mesure avec la photocopie; tous les piratages des ouvrages encore couverts par le droit d’auteur sont possibles.

BERTHO-LAVENIR, *Histoire des médias, de Diderot à Internet*, Paris, Armand Colin, 1996, pp. 324-236.

⁵⁹ Alain GIFFARD et Michel RICHARD, *op. cit.*, p. 12.

⁶⁰ André FRANÇON, *La propriété littéraire et artistique*, Paris, Presses universitaires de France, coll. “Que sais-je?”, 1979, p. 73.

⁶¹ Une explication de ces revendications serait l’augmentation du nombre de livres empruntés en bibliothèque par rapport au nombre de livres neufs vendus pour une période de un an. En 1995 on parle d’un rapport jamais atteint de 1 pour 3. Cf. Catherine BÉDARIDA, “Les bibliothécaires défendent la gratuité des prêts”, *Le Monde*, 21 avril 1995, p. 30.

⁶² Serges EYROLLES (entretien avec Élie Deschodt), “Dites-le avec des livres”, *Valeurs actuelles*, 14 mars 1994, pp. 62-63.

Qu'à cela ne tienne! Les responsables de la BNF, qui ne veulent déroger aux principes intellectuels d'exhaustivité et de cohérence de la collection numérisée, entament un processus de négociations difficiles avec les éditeurs. L'objectif est d'arriver à une entente dans le respect des règles du droit. Il s'agit de conclure sur des conditions de coûts "*supportables pour les utilisateurs*" – c'est-à-dire tels qu'ils ne dissuadent pas l'usage de la collection numérisée – et "*acceptables par les auteurs et les éditeurs*"⁶³.

Les négociations se déroulent de façon chaotique. En novembre 1991, la BNF, le Syndicat national de l'édition et des sociétés d'auteur signent deux protocoles d'entente. La BNF obtient le droit de numériser, dans les limites du droit sur la propriété littéraire, les documents qu'elle désire. Les ententes conclues avec une cinquantaine d'éditeurs sont limitées dans le temps: l'autorisation de numérisation vaut pour la phase d'expérimentation.

Les relations avec les éditeurs, en dépit de ces premiers protocoles, s'amenuisent, rendant l'avenir des textes numérisés toujours plus incertain. À l'automne 1994, le Syndicat national de l'édition, consterné par le fait que les négociations avec les éditeurs n'ont débouché sur aucune proposition de rémunération, demande aux éditeurs de refuser de négocier individuellement avec la BNF.

*Il n'est évidemment pas imaginable que la Bibliothèque nationale de France exploite les œuvres qu'elle a numérisées sans déterminer les conditions d'une rémunération effective des ayants droits, prenant en compte, sous une forme déterminée, la fréquence d'utilisation des ouvrages*⁶⁴.

À l'heure actuelle, les tarifs, les modalités de rémunération, le contrôle, etc., sont indéterminés. Mais "*s'il n'y a pas d'accord rapide, ce ne sont pas 100 000 fiches que l'on pourra consulter, mais un peu moins de 70 000!*"⁶⁵. La consultation des ouvrages récents, dont l'auteur est mort depuis moins de 70 ans, sera-t-elle même autorisée sur le PLAO? Ce n'est pas certain. En tout cas elle sera presque sûrement payante. Pour l'accès à distance, les éditeurs sont plus craintifs "*puisque l'usage (consultation ou reproduction) ne saurait être contrôlé par définition*". Mais il y aurait une solution intermédiaire: la possibilité de consulter à distance la collection à l'écran seulement, sans possibilité technique et économique de télécharger la totalité d'un ouvrage⁶⁶.

La BNF semble avoir joué avec le feu, en sortant de sa "niche" économique classique: la préservation du patrimoine. Les éditeurs ne se lassent pas de le lui rappeler, mais butent sur la puissance symbolique que représentent tant l'institution que la nouvelle technologie proposée. Le fléau de la justice penche cependant en leur faveur.

Toutefois, la BNF aurait pu présenter une collection numérisée cohérente et encyclopédique des ouvrages *du domaine public* ce qui aurait rempli à la fois sa mission patrimoniale et son goût pour la mission éditoriale.

⁶³ Cf. Marcelle BEAUDIQUEZ et Yannick MAIGNIEN, *La politique de numérisation de la Bibliothèque nationale de France*, 5-6 septembre 1994, 14 p.

⁶⁴ Serges EYROLLES cité par Laurence SANTANTONIOS, "Textes numérisés: les éditeurs font blocs", *Livres hebdo*, n° 136, 11 novembre 1994, pp. 21-23.

⁶⁵ André ZYSBERG (entretien avec Jean-François Augereau), *op. cit.*

⁶⁶ Cf. EUTELIS, *op. cit.*

F. — DIFFUSION ET PATRIMOINE, LE CHOC DES CULTURES

La mission de diffusion et de communication des ouvrages, que sert en partie une collection numérisée, bouleverse aussi l'institution elle-même: elle propose un nouvel équilibre dans les missions de la Nationale.

Depuis 1537, année de la création du dépôt légal⁶⁷ par François I^{er}, la rue de Richelieu a d'abord une vocation patrimoniale. C'est sa mission et sa raison d'être: conserver le patrimoine intellectuel français. La tradition de conservation du patrimoine est telle qu'elle assure à la BN une réputation internationale. "À Munich ou à New York, s'il faut identifier la main qui a enluminé un manuscrit du XIII^e siècle, on s'adressera à une poignée d'experts [...] rue de Richelieu"⁶⁸. Ses spécialistes, ses ateliers de restauration et de conservation sont réputés.

À l'inverse, les conditions d'accès à la rue de Richelieu provoquent l'indignation de lecteurs étrangers: tel étudiant américain s'étonne qu'on lui demande ses diplômes pour la consultation d'un livre; tel professeur à l'Université de Dublin regrette les procédures d'admission trop bureaucratiques qui l'empêchent de terminer sa recherche à temps⁶⁹. De fait, à la BN, la communication des documents est le privilège d'une élite d'universitaires et de gens de lettres (à l'opposé de la Bibliothèque du Congrès "ouverte à tous parce qu'elle appartient à tous"). Seuls 900 000 ouvrages sont communiqués, en dépit d'un fonds documentaire important, alors que les bibliothèques universitaires en prêtent 8 millions, et les bibliothèques municipales 80 millions!⁷⁰

Or voilà que le texte numérisé, issu d'une culture de diffusion de l'information, doit être implanté dans cette institution à vocation patrimoniale. Cela provoque des étincelles. Jusqu'en 1994, la BN refuse tout net de prêter un fonds de 60 000 documents pour la numérisation! Cette résistance sera "effacée" mais jamais résolue: ce n'est que parce que les deux institutions BN et ÉPBF fusionnent qu'elle disparaîtra.

Enfin, personne ne connaît encore les conditions d'accès aux documents numérisés de la collection. Y aura-t-il d'abord une autorisation à demander à la BNF ou bien seront-ils accessibles à tous les publics? Pour le moment, la structure existante semble dicter sa loi: "Il est probable qu'il faudra que le tarif de télécommunication d'un livre ou d'une revue soit dissuasif à l'égard du grand public"⁷¹. Encore une limite à la diffusion des textes numérisés!

Les postes de lecture sur place, eux, semblent en phase avec la tradition, malgré leur volonté de bouleverser la lecture traditionnelle. S'adressant exclusivement aux universitaires, au public actuel de la Bibliothèque nationale, le poste de lecture participe de l'*habitus*: l'outil est orienté pour "les lecteurs ayant des besoins impliquant de véritables techniques de lecture" auxquels la BNF veut en priorité répondre⁷². Ainsi, il prolonge la distinction physique des pratiques de lecture. Comme la BN, le PLAO a son public... en les murs.

⁶⁷ Tout ouvrage imprimé en France doit être déposé à la Bibliothèque nationale. Cette mission de conservation du patrimoine distingue la BN des autres bibliothèques françaises.

⁶⁸ Sylvaine PASQUIER, "Des trous dans la mémoire", *L'Express*, 4 mars 1988, p. 48.

⁶⁹ Cf. Christian BAUDELLOT et Claire VERRY, "Profession: lecteur?", *Bulletin des bibliothèques de France*, t. 39, n° 4, 1994, p.11.

⁷⁰ Patrice CAHART et Michel MELOT, *op. cit.*, p. 21.

⁷¹ Jean Favier, cité par L.C., "La Bibliothèque nationale de France ne veut pas "ruiner l'édition française"", *Le Figaro*, 15 décembre 1994.

⁷² Cf. BIBLIOTHÈQUE DE FRANCE, *Les postes de lecture assistée par ordinateur de la Bibliothèque de France*, *op. cit.*, p. 8.

Le match paraît joué d'avance. Le PLAO s'insère dans les structures existantes liées aux pratiques de lecture institutionnelles. Or, l'accès à distance à la collection propose de bouleverser ces pratiques en abolissant les distinctions physiques des pratiques de lecture. N'est-ce pas là une façon d'expliquer pourquoi la réalisation d'une collection numérisée associée au PLAO s'impose au détriment de l'accès à distance?

G. — DES LECTEURS TRÈS PARISIENS

Les lecteurs eux-mêmes, pourtant concernés par la mission de communication de la Bibliothèque, brillent par leur absence dans le débat. Ils avaient été pourtant plus "bavards" sur les aspects de l'architecture, de l'accueil des publics et de la répartition des collections entre la BN et la nouvelle Bibliothèque. L'introduction des nouvelles technologies de l'information n'anime ni de façon unanime ni avec le même intérêt la communauté intellectuelle des lecteurs de la BN. Les réactions semblent hétérogènes.

Les sceptiques rappellent que l'accès à distance et l'hyper-informatisation de la Bibliothèque reposent sur des incertitudes difficiles à éclairer⁷³, que le développement de la Bibliothèque s'éloigne des usagers qui ont des besoins "*modestement artisanaux*" pour un "*futurisme polytechnicien*"⁷⁴, que l'usage massif des technologies augmente le coût de fonctionnement et les dépenses⁷⁵, que le projet informatique est irréaliste et trop complexe et que l'EPBF privilégie les études sur la numérisation et sur le PLAO au détriment d'autres dossiers prioritaires⁷⁶.

D'autres, plus optimistes, voient dans les technologies de l'information et la numérisation les moyens d'assurer le rayonnement de la BNF en France et à l'étranger⁷⁷. Certains se réjouissent qu'en misant sur les nouvelles technologies de l'information, la BNF soit effectivement une bibliothèque d'un autre âge⁷⁸. D'autres enfin perçoivent l'engagement informatique comme un aspect positif de la BNF, un défi à relever, un pari sur d'autres supports que le livre⁷⁹.

Mais en dehors de ces réflexions isolées, le "lecteur λ " reste silencieux. Serait-ce par manque de connaissances techniques que les lecteurs interviennent peu sur l'enjeu de la numérisation et des technologies d'information en général? C'est probable. Un lecteur historien et non des moindres n'a-t-il pas mentionné en évoquant ces nouvelles technologies que "*(c)'est un projet ambitieux en ce sens que, comme on ne connaît pas la technique, il faut faire confiance aux techniciens*"⁸⁰. Une confiance qui laisse la voie libre aux "spécialistes" d'une lecture assistée par ordinateur qui n'existe pas encore. Indice: même dans une revue parisienne de débats où est accordée la priorité aux arguments des

⁷³ B.B., "Le rêve télématique", *Le Monde*, 14 février 1989, p. 2.

⁷⁴ Pierre NORA, "Bibliothèque de France: d'où venons-nous, où allons-nous?", *Le Débat*, n° 70, mai-août 1992, pp. 113-120.

⁷⁵ Francine MASSON, "Un laboratoire irremplaçable", *Le Monde*, 13 mars 1993.

⁷⁶ Jean-Pierre BARDET et Pierre CHAUNU, "Bibliothèque nationale – Éviter le pire?", *Commentaire*, printemps 1993, n° 61, vol. 16, pp. 113-119.

⁷⁷ Michel MELOT, "De la Bibliothèque nationale à la Bibliothèque de France", *Le Débat*, n° 55, mai-août 1989, pp. 137-142.

⁷⁸ Pierre NORA, "Dans le bon sens", *Le Débat*, n° 62, novembre-décembre 1990, pp. 4-8.

⁷⁹ Jean-Luc PINARD-LEGRY, "Bibliothèque de France, les chantiers de l'utopie", *Magazine littéraire*, n° 293, novembre 1991, pp. 98-104.

⁸⁰ André MIQUEL, "La Bibliothèque de France est un lourd investissement", *Réseaux*, n° 13, novembre-décembre, janvier 1993, p. 3.

futurs usagers de la Bibliothèque⁸¹, sitôt que l'on parle "technique", on laisse les experts trancher.

À cette concentration autour d'un noyau d'experts s'ajoute une concentration géographique: les lecteurs impliqués sont parisiens. Ce qu'on pourrait appeler le "lobby de la BN", constitué essentiellement des membres d'une élite parisienne, ne voit pas d'intérêt à l'accès à distance aux documents de la Bibliothèque. D'abord, les documents sont sur place. Ensuite, on leur sert sur un plateau un outil expérimental prestigieux: le poste de lecture. Enfin, le lobby a de quoi se réjouir: l'informatisation complète des fiches de la BN, l'intérêt majeur de la BNF pour la BN⁸², est réalisé.

CONCLUSION. — UN PARI DANS LE BROUILLARD

L'analyse de l'évolution du programme de numérisation a montré comment il s'était inscrit dans le contexte général de "perpétuelle" évolution des technologies de l'information, et dans le contexte particulier du projet présidentiel de la BNF. Nous avons pu montrer que les caractéristiques du programme sont le résultat d'un ensemble de choix⁸³, fondés sur la technique certes, mais aussi sur des principes et des idées, eux-mêmes inscrits dans une organisation qui a ses valeurs et son mode de fonctionnement. Avec une collection de 100 000 documents, la réalisation d'un poste de lecture assistée par ordinateur et le désir initial de rendre les documents télé-accessibles, le programme de numérisation est aussi ambitieux que l'était le projet présidentiel.

Par plusieurs aspects, le programme de numérisation relève du pari: un pari sur la transmission facile du mode "image", sur l'acceptation sociale d'une lecture assistée par ordinateur, sur un coût de fonctionnement acceptable⁸⁴, etc. D'aucuns proposent de ne pas "bouder" la "modernité qui galope", et ne pensent plus qu'à la mise du pari. Cela relève peut-être d'un optimisme irréfléchi, d'une volonté de "faire" l'Histoire à tout prix et, en définitive, d'un goût du danger peu compatible avec les exigences habituelles de l'État.

Quels sont les risques d'un tel pari? Que risque-t-on d'y gagner? Les incertitudes qui planent sur le programme de numérisation sont, en fait, de trois ordres: 1) technique, 2) économique, et 3) intellectuel et culturel.

Tout d'abord, des incertitudes techniques pèsent sur le contenu de la collection numérisée. Quelle formation faudra-t-il dispenser aux lecteurs de la BNF pour qu'ils se servent du PLAO et de la télétransmission? Comment rendre les outils informatiques si conviviaux et interactifs qu'ils seront utilisables par tous? Qui dit également que le lecteur à distance aura à sa disposition un équipement informatique suffisamment puissant pour travailler sur la collection numérisée? Enfin, que se passera-t-il si les réseaux de télécommunication évoluent moins vite que l'espèrent les concepteurs du programme de numérisation: le mode "image" sera-t-il condamné à être lent et coûteux (malgré les logiciels de compression de données), tandis que des modes "texte" ou mixtes (tels ceux

⁸¹ Exemple: "Métamorphoses du livre et de la lecture", pp. 131-171, *Le Débat*, septembre-octobre 1995.

⁸² Cf. Emmanuel LE ROY LADURIE (entretien avec Anne PONS), "Le Roy Ladurie: la bibliothèque idéale", *L'Express*, 15 avril 1993, pp. 80-83.

⁸³ Car "la nécessité [de la modernisation] ne s'impose pas au point d'évacuer les choix". Cf. Philippe ROQUEPLO, *Penser la technique. Pour une démocratie concrète*, Paris, Seuil, 1983.

⁸⁴ On connaît les difficultés d'autres institutions grandioses à Paris où la structure est si majestueuse que les budgets de fonctionnement suffisent à peine et obèrent par leur maigreur la réalisation du programme prévu.

utilisés actuellement sur Internet) deviendront la norme internationale? À moins d'un assouplissement de la politique du tout-mode-“image”, ne risque-t-on pas de reproduire le scénario catastrophe de la norme SÉCAM pour la télévision, objectivement meilleure mais devenue plus dispendieuse en raison de la petitesse du marché?

Aux incertitudes techniques s'ajoutent des contraintes de type économique et financier. Certes, dans la sphère des services publics, les considérations économiques ne sont pas prioritaires. Mais le passé est émaillé de ces éléphants blancs dont le budget de fonctionnement est si gros qu'il handicape la gestion courante (salles fermées, animations hors service, programmation au compte-gouttes). Le programme de numérisation de la BNF pourrait bien rencontrer de tels écueils: ainsi, le stockage en mode “image”, très gourmand en espace mémoire, ne risque-t-il pas de faire exploser les coûts de fonctionnement de la BNF, provoquant ou bien une inflation de l'enveloppe dédiée, ou bien l'amputation d'une partie du programme? De la même manière, la concurrence avec la consultation gratuite de la collection imprimée classique s'avère d'emblée féroce, et risque de tirer vers le bas les prix de la consultation de la collection sur le PLAO. Faudra-t-il prendre une mesure coercitive, comme rendre indisponibles les ouvrages dans leur version originale quand ils sont numérisés, pour créer un public captif au PLAO? Enfin, quel public voudra payer ce que les prêts entre bibliothèques permettent de réaliser à faibles coûts (faire venir un ouvrage de San Francisco à Vancouver coûte peu mais demande un délai de trois semaines; l'immédiateté de la télétransmission justifiera-t-elle un coût élevé, et lequel?).

Finalement, des incertitudes intellectuelles et culturelles – les plus nombreuses et les plus intéressantes – pèsent sur le programme de numérisation, et en particulier, sur le contenu de la collection numérisée. Minutieusement constituée dans une logique d'offre, les textes fondamentaux de cette collection ont été sélectionnés en fonction des critères de cohérence et d'encyclopédisme, on l'a vu, mais aussi de rareté, de demande, de lecture approfondie. *Qui trop embrasse, mal étreint*, dit le proverbe: à force d'avoir voulu plaire à tout le monde mais étant limitée par des contraintes de budget, la collection ne risque-t-elle de ne répondre que partiellement à la demande, c'est-à-dire à certains spécialistes et pas à d'autres, à certaines disciplines et pas à d'autres, etc. De la même manière, même si plusieurs commissions d'experts ont été constituées, comment peut-on être sûr que la collection soit en phase avec les utilisateurs, que le conservatisme se lie agréablement au progressisme intellectuel, que l'offre coïncide avec les besoins? Les lecteurs ne pourront en juger que lorsqu'ils auront accès aux services de la collection numérisée⁸⁵.

Sur les PLAO, les incertitudes d'ordre intellectuel ne sont pas moindres. Le pari pris sur le surgissement d'un nouveau mode de lecture, à peu près encore inconnu, grâce à un outil nouveau, est audacieux. Mais les résistances liées au passage d'une lecture sur papier à une lecture sur écran ne sont pas négligeables pour autant. Qui plus est, elles ne datent pas d'hier: Platon ne disait-il pas de la chose écrite qu'elle était aliénante?⁸⁶ Les philosophes de la BNF penseront peut-être la même chose de la lecture sur écran. En définitive, le pari intellectuel sur le PLAO a l'arrière-goût d'une idéologie du début des années quatre-vingt, et que l'on croyait dépassée: celle d'une “bureautique sans papier”,

⁸⁵ Un tel désajustement entre l'offre et la demande de documents numérisés s'est déjà manifesté. Une étude à l'Université Stetson – aux États-Unis – montre que parmi les demandes d'articles faites dans le cadre du prêt entre bibliothèque, seulement 3,3% des articles demandés sont disponibles dans les bases de données commerciales. Plusieurs facteurs expliquent cet écart entre l'offre commerciale et la demande académique: la période couverte par les banques de données est trop restreinte – pas de vieux périodiques –, la sélection de périodiques à caractère général est inappropriée aux besoins académiques spécialisés, etc. Cf. David EVERETT, *op. cit.* Cependant, la BNF cherche à éviter l'écueil de la demande en organisant sa collection numérisée: les gens devraient savoir *a priori* ce qui s'y trouve.

⁸⁶ Cf. Roger BAUTIER, *op. cit.*

d'une *paperless society*. Au même moment, le succès foudroyant du *fax* est venu nous en rappeler la vanité. Dans la société de l'information, "les techniques informatiques évoluent plus rapidement que nos comportements"⁸⁷.

Enfin, la transmission des documents subit elle aussi des contraintes intellectuelles et légales. En s'émançant de son rôle classique de dépôt légal et en proposant une collection basée sur un choix éditorial, la BNF entre en conflit direct avec les acteurs du monde de l'édition pour qui la BNF se transforme en concurrent commercial. À cause de ce problème actuel et concret du droit d'auteur, le nombre de documents numérisés actuellement consultables risque de se réduire à moins de 70 000. Plus de trente mille documents de la collection connaissent un sort incertain, soit 30 % de l'investissement en textes numérisés! Pourra-t-on cependant les consulter sur écran (sans possibilité de duplication) à la BNF; c'est probable puisqu'on peut déjà consulter les ouvrages sur papier, même s'ils ne sont pas tombés dans le domaine public? Mais pourra-t-on les consulter à domicile? Comment sera géré le risque de duplication sur disquette, disque dur ou CD-ROM? Imposera-t-on une taxe, comme sur les photocopieuses? Quel tarif l'utilisateur ou la BNF devra-t-il alors payer pour l'accès ou la diffusion de ces textes? L'avenir des 30 millions de francs investis mais litigieux (n'était-ce pourtant pas prévisible?) est tributaire de la réponse à ces questions, des résultats des négociations ou des évolutions techniques.

Bref, le programme de numérisation des imprimés de la BNF continue d'être le laboratoire d'idées que l'on avait voulu à sa création. Les incertitudes, tant techniques, légales, économiques qu'intellectuelles, sont à la mesure du risque encouru et des promesses du projet. Pour le moment, le programme est une suite d'inventions géniales. De nombreux facteurs interviendront pour en faire une innovation, dont le succès est par essence imprévisible et pour laquelle on ne s'étonne donc pas de trouver des incertitudes avant la "mise en public".

Espérons que l'ambitieuse logique d'offre se libère des pièges qu'elle s'est elle-même tendus et qu'en 1998, dix ans après le lancement du projet, les solutions "révolutionnaires" n'auront pas accouché d'une souris technologique. Ne serait-il pas troublant de constater que le programme de numérisation qui devait servir la démocratisation du savoir est devenu un système fermé générateur d'exclusion? Pis encore. De constater simplement que les usagers, pour diverses raisons, n'en veulent pas. Combien de ces incertitudes auraient pu être évitées?

⁸⁷ Shoshana ZUBOFF, "Nouveau contrat social informatique", *Pour la science*, n°217, novembre 1995, p. 170.