

La numérisation à la Bibliothèque nationale du Québec

Suzanne Ledoux et Philippe Martin

Volume 45, numéro 4, octobre–décembre 1999

Édition électronique

URI : <https://id.erudit.org/iderudit/1032723ar>

DOI : <https://doi.org/10.7202/1032723ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association pour l'avancement des sciences et des techniques de la documentation (ASTED)

ISSN

0315-2340 (imprimé)

2291-8949 (numérique)

[Découvrir la revue](#)

Citer cet article

Ledoux, S. & Martin, P. (1999). La numérisation à la Bibliothèque nationale du Québec. *Documentation et bibliothèques*, 45(4), 195–200.
<https://doi.org/10.7202/1032723ar>

Tous droits réservés © Association pour l'avancement des sciences et des techniques de la documentation (ASTED), 1999

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

érudit

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>

La numérisation à la Bibliothèque nationale du Québec

Suzanne Ledoux

Responsable de la base de données Iris
Bibliothèque nationale du Québec
s_ledoux@biblinat.gouv.qc.ca

Philippe Martin

Chef de division - Système informatique
Bibliothèque nationale du Québec
ph_martin@biblinat.gouv.qc.ca

En 1991, la Bibliothèque nationale du Québec (BNQ) faisait l'acquisition d'un système intégré de gestion documentaire. En 1993, la base de données bibliographiques Iris de la bibliothèque était accessible à distance par l'intermédiaire des réseaux Datapac et Internet au moyen de Telnet. En 1995, la compagnie Best-Seller, fournisseur du système intégré de la BNQ, présentait le prototype d'un catalogue (OPAC) en mode graphique basé sur la technologie Web laquelle permettait de rattacher des extensions multimédias (images, sons, textes) aux notices bibliographiques.

En 1994, le Fonds de l'autoroute de l'information (FAI) lançait son premier concours de création de contenus et de services en français dans Internet, avec financement à la clé. Best-Seller, en collaboration avec la BNQ, soumettait le projet *BEST WEB - Catalogue de bibliothèque multimédia*. Ce projet avait un double but : permettre à une compagnie québécoise d'être à l'avant-garde en lui permettant de développer le prototype de son catalogue multimédia, et permettre à la BNQ de numériser une partie importante de ses collections et de la diffuser dans Internet en annexe à son catalogue. Le montant total du projet s'élevait à 1 760 000\$ et la subvention possible à 450 000\$. Le projet était accepté en décembre 1995 et le FAI alloua une période de deux ans pour sa réalisation.

L'échéancier prévoyait la création d'un serveur Web dynamique (définition des fonctionnalités, analyse des paramètres, design des pages, développement du moteur de recherche, test, révision de l'interface), la définition des modalités de gestion des liens multimédias, des tests avec la base de données de la biblio-

thèque et avec quelques notices contenant des liens multimédias ainsi que la création des extensions multimédias (numérisation d'images, de sons, de textes, liens entre les fichiers numériques et les notices bibliographiques).

En juin 1996, à l'occasion du MIM (Marché international du multimédia), la BNQ a été en mesure, pour la première fois, de montrer Iris, un catalogue multimédia contenant quelques dizaines d'images et de sons.

Que faut-il numériser* ?

Il était primordial, avant de choisir les titres à numériser, de se demander pourquoi le faire. Pour la Nationale, l'attrait principal était d'assurer ainsi une meilleure diffusion du patrimoine documentaire québécois. Les collections spéciales ont été privilégiées afin d'en mieux faire connaître la richesse et d'en faciliter la consultation à distance. Ces documents étant, bien sûr, exclus du prêt entre bibliothèques, on ne pouvait en prendre connaissance qu'en les consultant sur place. On voulait aussi constituer un premier corpus québécois d'œuvres numérisées, un équivalent du Gallica de la Bibliothèque nationale de France. Si on peut lire Victor Hugo à l'écran, pourquoi pas Laure Conan ?

L'aspect conservation n'a pas été un critère de choix même si le document numérique constitue un support de substitution comme l'est la microfiche et, de ce fait, facilite la conservation des collections en diminuant la manipulation des originaux. En rendant plus facilement accessibles les documents iconographiques (estampes, affiches, cartes postales, etc.), ceux-ci seront probablement plus consul-

tés, mais le recours aux originaux sera limité. Par exemple, un usager peut rapidement visualiser une quarantaine de cartes postales de pont couverts, mais il choisira peut-être de ne consulter ensuite que cinq originaux. Il est peu probable que l'usager venu sur place choisisse de consulter la version numérique d'un document ; par contre, l'usager éloigné consultera la version numérique plutôt que de demander un prêt entre bibliothèques.

Sans faire un relevé chronologique des décisions qui ont mené au plan de numérisation, ce sont les documents iconographiques des collections spéciales qui se sont de prime abord imposés : estampes, affiches, livres d'artistes et cartes géographiques auxquelles se sont ajoutées les cartes postales, d'abord négligées parce que non traitées. L'intérêt manifesté pour les quelques extraits sonores rendus disponibles, dès le départ, a inspiré la numérisation de quelque 1 500 chansons. Enfin, comme à cette époque les premiers textes numérisés dans leur intégralité faisaient leur apparition dans Internet, il allait de soi qu'il fallait suivre le mouvement et numériser un ensemble de textes québécois. Compte tenu du budget dont disposait la bibliothèque et du coût de numérisation à l'unité, le plan de numérisation présenté au Tableau 1 a été adopté.

Certaines difficultés techniques ont amené une révision à la baisse du nombre de pages numérisées en mode image et surtout en mode texte et ainsi augmenter le nombre d'images numérisées.

Aperçu des réalisations

Côté estampes, la bibliothèque a opté pour une numérisation exhaustive de la collection. Plus de 9 000 estampes ont été

* Liste des acronymes à la fin du texte.

numérisées et sont accessibles ; il y en a actuellement 4 000 en cours de numérisation. Seulement une dizaine d'œuvres de très grand format ont été laissées de côté. Un choix de 1 500 affiches a été fait sur les 6 000 affiches signalées dans Iris. On a numérisé une sélection de 6 500 cartes postales. On a d'abord envisagé d'en numériser 20 000, mais le pauvre contenu documentaire de nombreuses cartes postales (ex. : un écureuil sur le Mont-Royal) et les sujets souvent redondants (plusieurs dizaines de cartes postales du pont Jacques-Cartier), nous ont obligés à faire un tri. Du côté des cartes géographiques, le choix s'est porté sur les 2 000 plans d'assurance-incendie de Montréal. La numérisation s'est faite à partir de la version microforme des originaux conservés à la British Library. Il a d'ailleurs fallu obtenir la permission de cette dernière, laquelle a cependant refusé la diffusion sur le Web. Mais les plans sont consultables sur place. On numérise actuellement 1 500 cartes des collections de la bibliothèque. En ce qui a trait aux enregistrements sonores, on a sélectionné, dans la collection 78 tours, 1 500 pièces musicales représentant le répertoire du début du siècle. Près de 500 livres d'artistes et de bibliophilie ont été choisis pour leur originalité, leur reliure et la qualité de leurs illustrations et des textes. Dans un domaine voisin, on a retenu la collection Massicotte constituée de 6 000 illustrations portant sur le Montréal des années 1870-1920. Très fragile, elle avait été microfilmée il y a quelques années. Enfin, 6 000 illustrations diverses tirées en grande partie de trois publications en série de la fin du siècle dernier (*Le Monde illustré*, *l'Album illustré* et *l'Opinion publique*) ont été intégrées à la collection numérique.

Considérations techniques

Aspects catalographiques

La création d'un catalogue multimédia nécessitait que les documents numérisés soient traités pour qu'on puisse y faire référence dans les notices bibliographiques, ce qui n'était pas le cas des affiches et des enregistrements sonores. Il a donc fallu chambarder quelque peu les priorités de traitement et procéder à la description

Illustration 1 : Le codage MARC qui permet en cliquant sur la vignette de basculer vers l'image plein écran

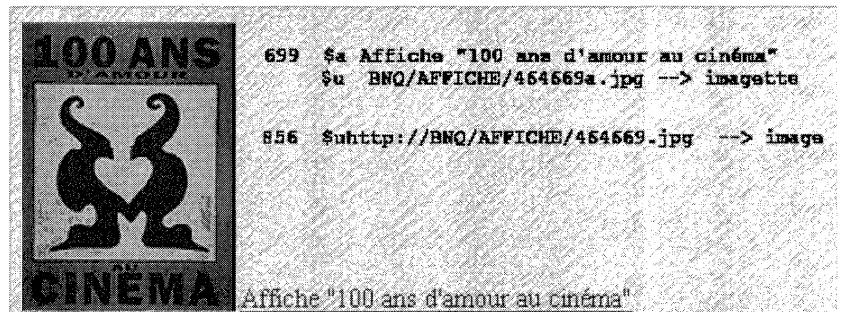


Tableau 1 : Plan de numérisation et réalisation

	Prévu	Réalisé
Estampes	8 500	9 000
Affiches	1 500	1 500
Cartes postales	5 000	6 500
Livres d'artistes	400	500
Cartes géographiques	2 000	2 000
Livres (mode image)	2 000	1 500
Livres (mode texte)	100	10
Pièces musicales	1 000	2 000
Illustrations de la Collection Massicotte	6 000	6 000

et à l'analyse de ces documents avant que ne débute leur numérisation. Les cartes postales constituaient un cas particulier : elles n'étaient pas traitées et il était impensable d'en traiter rapidement 5 000 en respectant les normes en vigueur du catalogue Iris ; un traitement minimal a été effectué dans une base de données indépendante à l'aide du logiciel *Inmagic*. C'est aussi cette solution qui a été privilégiée pour les quelque 6 000 illustrations tirées de revues anciennes.

Le traitement était indispensable afin de constituer un outil de travail permettant de gérer le processus depuis la sélection des documents jusqu'à leur numérisation et à leur intégration dans la banque Iris. En collaboration avec le fournisseur Kodak,

on a mis au point une base de données contenant pour chaque document les informations suivantes : [numéro séquentiel] - [cote du document] - [numéro de séquence du document] - [numéro du négatif] - [numéro du CD] - [position de l'image sur le CD] - [numéro du CD duplicata] - [auteur] - [titre] - [légende].

Au départ, certaines données sont extraites de la base Iris ou d'une base *Inmagic* et les autres champs sont remplis au fur et à mesure de l'avancement dans le processus. Des listes imprimées ont aussi été prévues pour gérer la sortie des documents des magasins et valider les prises de vue.

Le format MARC contenait depuis peu une étiquette relationnelle entre la notice et le fichier numérique, soit l'étiquette

856. Cette dernière contient l'adresse URL du fichier numérique. En collaboration avec Best-Seller, on a élaboré une étiquette locale pour annoncer, à l'aide d'une vignette et d'une légende, le contenu du fichier numérique.

Numérisation d'images

En 1996, au point de départ du programme, les techniques utilisées n'étaient pas les mêmes que maintenant. Les caméras numériques de type professionnel n'avaient pas encore atteint leur qualité actuelle, principalement dans le rendu des couleurs. Aussi a-t-on opté pour une approche plus classique, mais qui assurait de meilleurs résultats : prise de vue en couleurs, numérisation du négatif, report sur PhotoCD en cinq résolutions (procédé Kodak) et production de l'image plein écran et de la vignette pour les pages Web.

Les premières images numérisées furent celles des estampes et, lors des premiers essais, c'est le film Ektachrome 64, en 35 mm qui donna les meilleurs résultats. À la vérification des premiers PhotoCD, on découvrit des points blancs sur les images, lesquels se sont révélés être des poussières provenant de l'électricité statique engendrée par la manipulation de la pellicule. Le jet d'air de 15 livres destiné à nettoyer le film n'était pas suffisant et un jet plus puissant risquait de briser la pellicule. Il a donc fallu opter pour la diapositive montée dans un cadre de carton pouvant, elle, supporter 35 livres de pression d'air au nettoyage. Pour ne pas reprendre les 2 000 premiers clichés, Kodak a décidé de monter ces films en diapositives et de renumériser. Lors de la contre-vérification, un autre problème a surgi : certaines images étaient tronquées, l'appareil coupait le film pour en faire des diapositives et se désajustait insensiblement d'une image à l'autre. Certaines photos ont donc dû être reprises.

On a aussi détecté des images mal centrées. Kodak devait régler ce problème provenant du mouvement mécanique du chargeur qui faisait bouger les diapositives dans leurs cadres et les décentrait. En conséquence, ce n'est qu'après quelques semaines de vérification et de contre-vérification que le programme a atteint sa vitesse de croisière.

Comme support d'archivage des images numérisées, on a retenu le PhotoCD de Kodak qui propose cinq résolutions à partir de l'image produite lors de la numéri-

sation et permet de disposer d'autres tailles aux fins de reproduction à plus petite ou à plus grande échelle. Il est à noter que la taille maximale ne peut dépasser 18 Mo (sauf dans le PhotoCD-Pro où elle peut atteindre 72 Mo). Un PhotoCD contient une centaine d'images, c'est dire que les 9 000 estampes représentent 900 PhotoCD.

Cependant, le PhotoCD est inapproprié pour des images de grand poids comme les cartes géographiques. Pour ces dernières, on doit – si l'on veut obtenir une bonne lisibilité à l'écran – constituer une mosaïque à partir de l'image numérisée, en divisant cette image en deux ou quatre parties (ce qui nécessite un travail de manipulation à l'aide d'un logiciel d'édition d'images, comme *PaintShop Pro*). Pour ces images de très grande taille, un archivage sur cédérom standard est approprié après conversion en format JPEG compressé au dixième du format de numérisation.

À partir des types de supports d'archivage (PhotoCD et cédérom), on a créé deux séries d'images pour les pages Web, toutes deux au format JPEG : une vignette d'environ 12 Ko et une grande image d'environ 120 Ko.

Numérisation de textes

On a opté pour une numérisation de masse en mode image (le mode texte appliqué à quelques documents l'a été à titre expérimental) et on a adopté PDF d'Adobe comme format de diffusion. Ce format permet de présenter à l'écran le document tel qu'il est dans sa forme originale. Il est visualisé facilement et gratuitement sur toutes les plates-formes.

L'inconvénient majeur est la taille excessive de certains fichiers. Par exemple, le fichier du premier tome du *Dictionnaire généalogique des familles canadiennes depuis la fondation de la colonie jusqu'à nos jours*, de l'abbé Cyprien Tanguay fait plus de 100 Mo. On a préféré garder intact la chaîne de numérisation et régler dans une seconde étape les cas les plus flagrants en fractionnant les fichiers PDF.

Malgré l'utilisation d'un numériseur à plateaux mobiles, quelques documents ont été mis de côté en raison d'une reliure trop serrée ou de marges trop étroites. À la vérification, certains documents se sont avérés illisibles à cause d'un effet de transparence d'un papier trop mince. Les pages

principales (page de titre, pages liminaires, index et table des matières) ont été systématiquement vérifiées pour chaque document, ainsi qu'un certain nombre de pages aléatoirement choisies afin de détecter des sauts de pages, des pages mal cadrées, des pages illisibles ou des pages incomplètes. Lorsque le document est jugé satisfaisant, il est archivé sur un cédérom standard. L'archivage comprend les fichiers TIFF, le fichier PDF et dans quelques cas un fichier ACD lorsqu'il y a reconnaissance de caractères.

Numérisation de son

Près de 2 000 faces de disques 78 tours ont été enregistrées sur rubans DAT au moyen d'une enregistreuse numérique, directement à partir des disques après leur dépoussiérage, puis reportées sur des cédéroms standards. De plus, leur contenu a été traité informatiquement à l'aide d'un logiciel de « nettoyage » pour éliminer les bruits parasites et restituer les signaux sonores endommagés avant d'être convertis au format WAV.

Trois minutes de voix avec accompagnement orchestral échantillonnées à 22 KHz en 16 bits donnent un fichier d'environ 7 Mo. Une fois compressé au format RealAudio, nous obtenons un fichier d'environ 700 Ko. Il est à noter que les codeurs RealAudio actuels permettent de garder dans un seul fichier plusieurs taux de compression qui sont déchargés dans Internet en fonction de la capacité du lien de connexion de l'utilisateur.

Cahier des charges

Il ne faut pas oublier l'importance de bien rédiger son cahier des charges avant d'entreprendre tout programme de numérisation et de bien préciser : le type, le format, l'état physique et la quantité de documents à numériser ; les étapes souhaitées du processus de numérisation ; les sous-produits attendus ; le format des fichiers et leur taux de compression ; les supports ; les contrôles de qualité et les assurances couvrant la détérioration des documents consécutive à la manipulation.

L'expérience a montré qu'il est préférable de retenir – si les procédures administratives le permettent – un fournisseur reconnu, même si ses prix sont un peu plus élevés. Il peut également être avantageux de regrouper les documents à numé-

riser en lots de même type et d'autoriser les soumissions sur un ou plusieurs lots pour un même fournisseur. Il ne faut pas oublier non plus d'indiquer dans les clauses de l'appel d'offre que le contrat du fournisseur retenu ne sera signé qu'après les tests par le fournisseur sur un échantillon de documents et qu'après l'acceptation de ces tests par le client.

Il est préférable que les prises de vue – qu'elles soient par caméra conventionnelle ou numérique – s'effectuent dans les locaux du client, mais que le fournisseur ait la responsabilité de l'aménagement du studio et de son équipement. Il est conseillé pour les formats 1 x 2 mètres d'utiliser un tableau à succion pour maintenir les documents en place sans aide. Pour les très grands formats, il faut envisager, pour le studio, un local à haut plafond (4 mètres) afin d'être à même de photographier des documents posés sur le sol.

Chaîne de travail

Pour tous les types de documents à numériser, la chaîne de travail est sensiblement la même – chez le client tout au moins. Mentionnons les étapes habituelles : la préparation des documents, c'est-à-dire leur choix, leur traitement bibliographique, s'il n'est pas déjà fait, leur manipulation et leur identification au moyen de listes de contrôle (numéro de document, de prise de vue, d'image numérisée, de support d'archivage et de position, numéro d'image et de vignette dans les pages Web) ; le processus de numérisation, c'est-à-dire les prises de vue, la numérisation elle-même, les corrections à l'écran de contrôle et la saisie des légendes ; la comparaison avec le document original, la production des supports d'archivage, le rangement des documents, la conversion des formats destinés au site Web, l'intégration des images réalisées dans le serveur et la création des liens catalogographiques.

Une fois les fichiers JPEG disponibles sur le serveur, des liens doivent être établis entre eux et les notices bibliographiques. C'est l'étiquette MARC 856 qui effectue le lien entre la notice et le fichier numérique, tandis que l'étiquette locale

656 \$a Légende qui s'affiche à côté de la vignette
\$u Adresse URL de la vignette
856 \$u Adresse URL du fichier numérique

699 gère cet accès à l'aide d'une vignette intégrée à la notice.

Dans Iris, l'adresse URL est associée à deux tables : une première définit le chemin d'accès du serveur Web où sont situés les fichiers numériques et une deuxième définit le répertoire.

Ex. : Su <http://BNQ/AFFICHE/464669.jpg>
BNQ correspond à l'adresse du serveur Web pour les fichiers numériques :
www2.biblinat.gouv.qc.ca
AFFICHE correspond au répertoire qui contient les fichiers JPEG des affiches. Si on réorganise les données sur un autre serveur ou dans une nouvelle structure de répertoire, il n'y a qu'à changer les informations dans la table et non dans les notices.

L'insertion des étiquettes 699 et 856 dans chacune des notices se fait en lot à partir d'un fichier plat contenant les informations appropriées. Le fichier est produit à partir de la base de données.

Processus de numérisation

Le processus de numérisation varie selon que l'on a affaire à des estampes, à des affiches ou à des cartes postales, à des livres de bibliophilie ou à des cartes géographiques.

L'exigence principale lors de la numérisation d'estampes, d'affiches ou de cartes postales est la fidélité à l'original quant aux couleurs, aux contrastes et à la définition (« piqué »).

Les contraintes à respecter sont : la taille des fichiers après numérisation et le respect des droits d'auteur. Ce dernier aspect peut être contourné en ne diffusant que des vignettes sur le site Web, en ne montrant que de grandes images de qualité dégradée afin de les rendre impropres à une reproduction commerciale ou en leur apposant un filigrane ou un tatouage.

On doit respecter un certain nombre d'étapes :

1) la prise de vue, à plat ou verticalement selon les formats, et avec un éclairage direct – ou rasant dans le cas des gaufrures – en tenant compte du fait que certaines estampes peuvent être constituées d'un matériau bril-

lant ou réfléchissant. Avant chaque série de prises de vue identiques, il est recommandé de photographier un tableau des couleurs qui servira d'étalon lors de la vérification après numérisation ;

- 2) la vérification des négatifs ;
- 3) la numérisation proprement

dite au cours de laquelle aucune correction n'est apportée aux couleurs : ce sont les négatifs qui doivent être repris s'ils ne sont pas le plus fidèle possible aux originaux ;

4) la gravure des PhotoCD avec cinq résolutions ;

5) enfin, la création des vignettes et des

grandes images en format JPEG ; le taux de compression retenu doit représenter un bon compromis entre la taille de l'image et la fidélité à l'original.

En gros, les coûts, à l'issue du processus décrit ci-dessus, sont de l'ordre de 7 \$ par image. Et dernier détail, les visionneuses utilisées, après calibrage du moniteur sur lequel se fait la vérification, ont été Magic Lantern et ACD See pour un visionnement rapide et PaintSho Pro pour une vérification plus poussée.

En ce qui a trait aux livres d'artistes, aux livres-objets et aux ouvrages de bibliophilie, quelques éléments sont à souligner. La variété des supports et le fait que la troisième dimension est un facteur essentiel exigent un éclairage et un angle de prise de vue spéciaux, adaptés à chaque document.

Les diapositives des reliures, des couvertures, des pages constituées de textes ou d'images et des objets (blocs, rubans, coffrets...) sont vérifiées et reclassées avant d'être envoyées à la numérisation. L'ordre dans lequel les photos sont prises n'est pas nécessairement l'ordre de présentation privilégié.

Côté cartes géographiques, on a jugé convenable, après essais, de procéder à la numérisation directe des cartes et plans géographiques sur un numériseur d'environ un mètre de largeur : la longueur du document (par exemple, un rouleau) importe peu ; si la plus petite dimension de la carte dépasse un mètre, elle est numérisée en deux phases. Le coût du processus est d'environ 10 \$ par carte. Le coût pour créer une mosaïque à partir d'une grande

carte est établi sur une base horaire.

Qu'en est-il des pièces musicales ? Une fois le choix des pièces effectué et les disques dépoussiérés, l'enregistrement s'est fait sur une machine de type DAT, sans aucune correction de manière à ne pas altérer le signal. Cet enregistrement s'est fait en continu en écoutant chaque pièce pour être certain de sa bonne qualité sonore. Puis avec un logiciel de traitement du son (*Steinberg Sound processor & Denoiser*) chaque pièce est nettoyée de ses parasites sonores, compressée et convertie au format WAV. Cette étape permet aussi d'identifier les morceaux (numéro de correspondance avec les notices de la base de données, titre, interprète). Le coût du processus est d'environ 13 \$ par pièce.

Reste les monographies. Pour numériser des textes pris dans des volumes, il n'est pas nécessaire de passer par l'étape « prise de vue » si le format des pages du volume numérisé ne dépasse pas 30 x 45 mm. On peut utiliser un numériseur à plat avec alimentation feuille à feuille si le livre peut être massicoté ; sinon il faut utiliser un numériseur spécial (son prix peut atteindre 85 000\$ US!).

Le processus à suivre est le suivant :

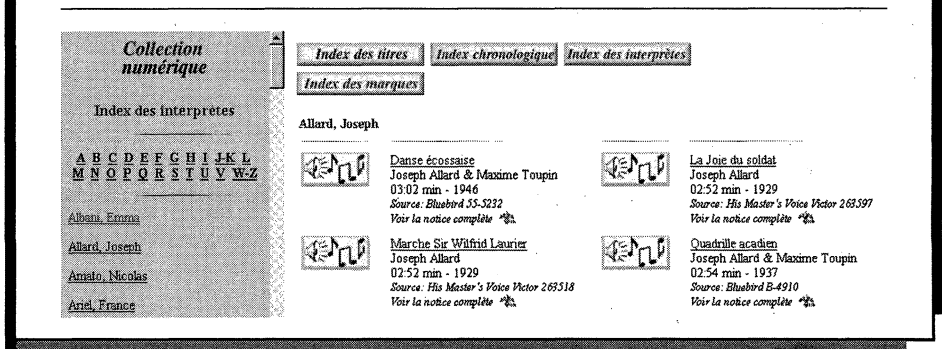
1. choix des documents, en fonction de leur rareté, de leur intérêt et de leur bon état physique ;
2. établissement d'un bordereau de numérisation ;
3. numérisation en mode « image » ;
4. traitement du fichier résultant par le logiciel *Acrobat* ;
5. traitement par un logiciel de reconnaissance de caractères (OCR) pour les documents que l'on veut indexer plein texte. Cela suppose : traitement, vérification, saisie manuelle du texte erroné, conversion en format PDF ; et enfin, archivage sur cédérom.

La visionneuse utilisée pour la vérification a été *Acrobat Reader* version 3.0. Le processus de numérisation revient à 0,30\$ par page en mode « image » et à 3,00\$ en mode « texte ».

Environnement WEB

En plus de donner accès à la collection numérique de la Nationale par le catalogue de la bibliothèque, nous avons développé pour chacun de ses volets un accès différent dans un environnement Web, afin

Illustration 2 : L'environnement Web des enregistrements sonores



de tenir compte de la spécificité de chaque catégorie de documents.

Les estampes, par exemple, sont regroupées par artistes et présentées en ordre chronologique de façon à ce qu'on puisse saisir d'un coup d'œil l'évolution picturale d'un artiste. Cet environnement permet aux historiens de l'art, aux professeurs et aux artistes eux-mêmes d'utiliser la collection à distance. L'environnement créé pour les affiches privilégie une organisation par grands sujets (cinéma, danse, livre, opéra...) et rend compte de l'évolution du domaine. Les notices minimales du corpus des monographies numérisées sont accessibles par les index traditionnels (auteurs, titres, sujets). On peut aussi, à partir de cet environnement, faire afficher la notice complète d'Iris. Cette fonctionnalité est aussi disponible pour les chansons numérisées. Dans ce cas, l'environnement Web met en évidence la chanson tandis que la notice Iris décrit le disque 78 tours. Aux index des interprètes, des sujets, des titres et des dates, on a ajouté un index des marques de fabricants, pour les spécialistes et les collectionneurs, lequel permet de repérer les chansons numérisées par le nom de la compagnie, par exemple RCA Victor, et par le numéro du disque.

Pour les estampes, les affiches, les textes et les chansons numérisées, les informations affichées et les clés d'accès sont tirées du catalogue Iris. Les livres d'artistes, quant à eux, présentent certaines particularités : il y a deux notices associées à chaque livre d'artistes. La notice Iris décrit le document selon les *RCAA2* ; une autre notice, beaucoup plus détaillée est destinée au catalogue des livres d'artistes publié par la BNQ. Le livre d'artistes numérisé est constitué d'images JPEG

accessibles depuis une série de vignettes. On peut afficher cette série à partir de la notice Iris ; dans l'environnement Web, elle est associée à la notice destinée au catalogue des livres d'artistes. Les points d'accès y sont ceux de la base de données établie pour produire le catalogue et gérer cette collection. On a donc une table alphabétique des artistes, des écrivains, des éditeurs, des relieurs et des titres.

Les autres volets de la collection numérique ne sont pas accessibles par l'intermédiaire du catalogue Iris : il y a d'abord les cartes postales repérables à l'aide du moteur de recherche *BestWeb Pro* et repérables par des index de lieux, de rues, de sujets, de noms et de titres. L'environnement Web développé pour les illustrations de la *Collection Massicotte* a comme point de départ un outil de travail préparé il y a quelques années sur traitement de texte et qui indexe les images de cette collection par noms de rues et par sujets. Les bibliothécaires de la référence ont insisté pour numériser 6 000 illustrations tirées du *Monde illustré*, de *l'Album illustré* et de *l'Opinion publique*. La richesse iconographique de ces illustrations sera mise en valeur par des tables nominatives et thématiques.

Conclusion

La numérisation a constitué et constitue encore un défi de taille, car bien que réalisée par des entreprises commerciales, une bonne partie du processus repose sur le personnel de la bibliothèque. Au moment de la planification des étapes de réalisation, nous avons prévu que le personnel participerait au choix des documents, à la manipulation des collections, à

la vérification et à la diffusion électronique.

L'étape de la vérification a été sous-évaluée. Nous pensions qu'il suffirait d'exercer un contrôle sévère au début et, par la suite, de vérifier par échantillonnage. De fait, toutes les images ont été vérifiées une à une. Ce n'est que pour les textes que la vérification s'est faite par échantillonnage. Faire le suivi des erreurs, les corriger et corriger les erreurs des corrections ont été les aspects les plus difficiles autant pour le personnel de la bibliothèque que pour les fournisseurs. Par contre, la participation du personnel, particulièrement celui des collections spéciales, à la mise en place de chacun des environnements Web s'est révélée gratifiante, bien qu'encore là, il a fallu normaliser, tester et corriger. Le personnel s'est pris au jeu et a proposé des outils de navigation de plus en plus complexes à réaliser. Ces efforts sont récompensés par la réaction du public et des chercheurs qui consultent notre site et nous font part de leurs réactions par courrier électronique. Les statistiques de consultation de notre

site Web sont encourageantes. Il y a deux ans, nous comptions une moyenne de 250 sessions par jour, chiffre qui atteint maintenant près de 700, et cela sans compter les visiteurs que nous ne pouvons comptabiliser, ceux qui se branchent directement dans le catalogue Iris sans passer par nos pages.

La disponibilité du catalogue en mode graphique dans Internet ainsi que son site Web ont augmenté considérablement la visibilité de la Bibliothèque nationale au Québec et partout dans le monde. La bibliothèque a d'ailleurs mérité un prix pour son catalogue lors du concours « The Best Library-Related Web Site » organisé par la revue *Library Hi Tech* en 1997.

Nous espérons, bien sûr, poursuivre la numérisation de nos collections et contribuer ainsi à la diffusion du patrimoine documentaire québécois. La collection numérique devra s'enrichir des documents nouvellement acquis et se développer en exploitant d'autres avenues; il y a d'ailleurs un projet en cours concernant les illustrations des monographies du XIX^e siècle.

Liste des acronymes

- ACD : fichier contenant les informations obtenues de la reconnaissance optique de caractères
- DAT : Digital Audio Tape
- JPEG : Format d'image compressée reconnu par les navigateurs Web
- PDF : Portable Document Format. Format de fichier créé par Adobe permettant de visualiser et d'imprimer un fichier sur n'importe quelle plateforme via Acrobat Reader
- TIFF : Tagged Image File Format
- WAV : Fichier audio sous Windows



289, Boulevard Desaulniers
Saint-Lambert, Québec
Canada J4P 1M8

Téléphone: (450) 671-3888
Télécopieur: (450) 671-2121

Accueillir, orienter, informer

Bertrand Calenge 69,95 \$

Action culturelle en bibliothèque

Viviane Cabannes et Martine Poulain **NOUVEAUTE** 63,95 \$

Concevoir, réaliser et organiser une bibliothèque

Danielle Taesh-Wahlen 57,00 \$

Conduire une politique documentaire

Bertrand Calenge **NOUVEAUTE** 72,95 \$

Droit d'auteur et l'édition

Emmanuel Pierrat 69,95 \$

Ouvrages et volumes

Anne-Marie Bertrand et Anne Kupiec 63,00 \$

Le Patrimoine

Jean-Paul Oddos **NOUVEAUTE** 69,95 \$