

LA NUMERISATION DES DOCUMENTS CARTOGRAPHIQUES ANCIENS: SUPPORTS TRADITIONNELS ET NOUVELLES TECHNOLOGIES*

By Pierre-Yves Duchemin

Résumé: La numérisation de documents graphiques, et notamment de documents cartographiques, est depuis longtemps un sujet d'étude à la Bibliothèque nationale. Dès sa création en 1994, la Bibliothèque nationale de France a prévu d'offrir à ses utilisateurs, un grand nombre d'images numérisées à partir de ses collections patrimoniales : elle s'adresse aux chercheurs, aux enseignants et aux étudiants et vise également un public plus large car les collections importantes de documents cartographiques ne sont pas nombreuses en France. Cette communication, centrée sur les documents cartographiques, aborde les divers problèmes qui se sont posés pendant la réalisation du projet.

Par ailleurs, la Bibliothèque nationale de France a produit des documents édités sur CD-Roms : « Mappamondi : une carte du monde au XIVe siècle » qui est une représentation numérique de l'Atlas Catalan, « Les globes de Louis XIV : la terre et le ciel par Vincenzo Coronelli », « La Géographie d'Idrîsî : un atlas du monde au XIIe siècle ».



Mappemonde dans un bonnet de fou (d'après Ortelius, ca 1580)

1. Historique des projets de numérisation d'images à la Bibliothèque nationale de France

La numérisation de documents graphiques, et notamment de documents cartographiques, est depuis longtemps un sujet d'étude à la Bibliothèque

* Paper presented at the 67th IFLA General Conference, Boston USA, August, 16-25 2001. Section on Geography and Map Libraries.

nationale : en 1989, fut produit un vidéodisque analogique sur la « Révolution Française » qui contenait des documents cartographiques de la fin du XVIII^e siècle. En 1991, un partenariat avec une société californienne pour numériser la collection de 437 portulans conservés au département des Cartes et plans et les éditer sur CD-Rom a été sur le point d’aboutir mais n’a pu voir le jour en raison de la loi française très stricte en ce qui concerne la propriété intellectuelle et l’usage du patrimoine national : le ministère de la Culture, tutelle administrative de la Bibliothèque nationale, n’a pas accepté cette procédure.



Portulan de Christophe Colomb (1492)

Dès 1994, la Bibliothèque nationale de France a prévu d’offrir à ses utilisateurs, dans un premier temps physiquement présents dans ses salles de lecture, dans un second temps distants, un grand nombre d’images numérisées à partir de ses collections patrimoniales ; en offrant la consultation d’images numériques réalisées à partir d’une partie de ses collections, la Bibliothèque nationale de France s’adresse évidemment à son public naturel, c’est-à-dire aux chercheurs, aux enseignants et aux étudiants de second et troisième cycle. Mais elle vise également un public plus large, bien connu depuis longtemps sur le site Richelieu, où sont conservées toutes les collections spécialisées ouvertes aux chercheurs, aux étudiants... et au « grand public », car les collections importantes de documents spécialisés, et notamment cartographiques, ne sont pas nombreuses en France.

Bien que ce programme ne soit pas destiné à une catégorie précise du « public à distance » de la Bibliothèque nationale de France, il n’en favorise pas moins, de par sa nature et ses caractéristiques, certains usages auxquels l’offre numérique existante n’apporte pas toujours une réponse satisfaisante : curiosité ponctuelle quant à la nature des collections, découverte plus détaillée de leur diversité, connaissance plus approfondie d’une catégorie de documents ou d’une pièce

précise – sans que cela soit exclusif de la satisfaction de demandes plus spécialisées ou plus érudites, qu'autorise la reproduction d'un certain nombre de pièces dans leur intégralité. L'insertion de ce programme dans l'offre numérique doit tenir compte de cette dimension d'« introduction aux collections » par l'établissement de liens avec les différentes composantes de cette offre numérique et certaines parties du site (Présentation des départements).

La variété des documents conservés sur le site Richelieu, ainsi que l'importance de la volumétrie des collections à traiter - on approche les 20 000 000 documents - montre combien il est difficile de concevoir une sélection et de maintenir une politique cohérente à l'intérieur de l'institution quand on doit traiter des documents aussi nombreux et aussi différents dans le cadre d'un projet de numérisation.

Bien qu'à la Bibliothèque nationale de France, la géographie ne soit pas considérée comme un champ spécifique de la connaissance - elle est traitée comme une discipline transverse et les documents cartographiques sont disséminés dans plusieurs départements, même si le département des Cartes et plans est le plus gros réservoir de représentations du monde - les collections retenues montrent une cohérence intellectuelle interne et sont prévues pour avoir une existence propre, avec des buts scientifiques et / ou artistiques, en tant qu'« unités documentaires ». Cette méthode permet d'éviter l'inconvénient qui consiste à offrir des ensembles thématiques incomplets.

2. Critères de sélection

Parmi les critères qui ont présidé au choix des collections, les images numériques ont été créées pour :

- permettre la préservation et la conservation du document original, notamment en ce qui concerne les documents de grand format ou les documents très fréquemment consultés,
- permettre la mise en valeur de collections en les ouvrant à un plus large public, grâce notamment à la transmission à distance,
- favoriser la mise en valeur de collections patrimoniales et prestigieuses de documents rares à valeur documentaire ou esthétique, souvent non publiés, en publiant des disques optiques compacts ou des Cd-webs,
- améliorer les possibilités d'accès au document en proposant un substitut électronique à l'utilisateur final,
- permettre un accès multiple au document,
- autoriser une communication plus importante et plus savante du document grâce à l'utilisation de stations de travail spécifiques (poste d'accès à la

bibliothèque numérique = PABN) qui permettent à l'utilisateur d'effectuer des zooms et d'obtenir une finesse de détails qu'il ne pourrait espérer à partir de l'original,

- offrir à l'utilisateur final une copie électronique du document sur un outil de travail spécifique dans un nouvel environnement technique,
- permettre des sorties payantes sur imprimante de qualité, à l'intérieur de l'établissement,
- un dernier critère a été la complétude et la cohérence de la reproduction photographique, ainsi que la présence de notices bibliographiques informatisées, sur fiches ou sur catalogues imprimés.

Enfin, en France, la situation juridique sur l'utilisation des images numériques n'est pas simple : les lois de propriété et de protection intellectuelle sont très strictes et on doit toujours rechercher d'éventuels ayants-droits. C'est la raison pour laquelle les projets d'images numérisées de documents cartographiques concernent des documents anciens, pour lesquels n'existent pas de problèmes de communication ou de reproduction, puisqu'ils appartiennent à la Bibliothèque nationale de France depuis des années, voire des siècles.

3. Projets d'images numériques de documents cartographiques

La première opération de numérisation d'images a consisté en la réalisation de 260 000 images de « non-livres », parmi lesquelles plusieurs dizaines de milliers de documents cartographiques, de plans d'architecture, de photographies à caractère ethnologique, etc.

Parmi les 11 projets mis en œuvre, 7 sont relatifs à des documents cartographiques ou nécessitent une indexation géographique : cartes de France, d'Europe, du Monde et du ciel des XVII^e et XVIII^e siècles de la collection d'Anville, plans d'architecture du XVIII^e siècle dessinés par Boullée et Lequeu, dessins et gravures de villes et de paysages français du XVI^e au XIX^e siècle de la collection Destailleur, photographies en noir et blanc de la Société de géographie prises en Afrique par les explorateurs français au XIX^e siècle, photographies de Paris en noir et blanc prises par Eugène Atget, photographies en couleurs de villes et de paysages français prises au cours des années 1980 pendant les missions DATAR, etc. Cette première série d'images numérisées est centrée sur la France, mais les nouveaux projets à l'étude possèdent une couverture géographique plus importante.

Les projets réalisés à la Bibliothèque nationale de France concernent des documents appartenant à :

- département des Cartes et plans (1 400 000 documents) qui regroupe cartes manuscrites, gravées et imprimées, atlas, globes, objets géographiques, dessins et notes d'explorateurs, photographies, plaques de verre, etc., de la plus ancienne carte connue à l'imagerie satellitaire, ainsi que les collections de la Société de géographie (livres, périodiques, cartes, photographies, plaques de verre, etc.).

Parmi ces collections, le choix s'est porté sur la prestigieuse Collection d'Anville (1697-1782) :

Parmi les 10 500 documents, dont un grand nombre d'*unica*, qui composent la collection réunie par Jean-Baptiste Bourguignon d'Anville au XVIII^e siècle et qui couvrent la terre entière, le projet de numérisation concerne 2 408 documents cartographiques libres de droits représentant la France, les provinces, les généralités, les diocèses, etc. du XVI^e au XVIII^e siècle, c'est-à-dire globalement l'organisation administrative, politique et religieuse de la France sous l'Ancien Régime, quelques cartes d'Europe, quelques mappemondes et quelques cartes du ciel. Est également incluse la première couverture régulière de la France réalisée à partir de 1747 par les Cassini, dont une version coloriée à la main en toises ayant appartenu à la reine Marie-Antoinette et une version noir et blanc métrique, pour laquelle un taux de résolution plus élevé a été nécessaire (4 000 x 6 000 points par pouce, soit 180 fichiers de 90 Mo, si l'on utilise la version couleur. La carte de Cassini pourrait faire l'objet d'une édition dans la collection « Sources » ; ce document représente en effet une mine toponymique pour la France de l'Ancien Régime. La numérisation des 4 346 images de la collection d'Anville (cartes en plusieurs feuilles) a été effectuée à partir de cibachromes 13 x 18 cm (microfiches monovues). Le catalogage a été effectué à la feuille dans BN-Opaline et certaines notices ont été regroupées en lots cohérents (cartes des diocèses, cartes des provinces, carte de France de Cassini, Gaule, assemblage de 5 à 6 feuilles représentant un cours d'eau, etc.).



Mappemonde cordiforme d'Oronce Fine (1536)



Feuille n°99 (Grenoble) de la carte de Cassini coloriée

- département des Estampes et de la photographie : plusieurs millions de documents, parmi lesquels on dénombre cartes, panoramas, plans d'architecture, etc.

Le choix s'est porté notamment sur les collections de documents produits par les deux architectes visionnaires du XVIII^e siècle Boullée et Lequeu, qui n'ont jamais dépassé le stade du dessin architectural.

- Dessins d'architecture d'Étienne-Louis Boullée (1728-1799)

120 pièces (dessin ou lavis) de grand format (150 x 50 cm) dont quelques documents concernant la Bibliothèque royale, notamment l'installation d'une verrière au dessus de la Cour d'honneur pour en faire la Salle de lecture. Quantitativement peu important, ce projet est néanmoins d'un intérêt considérable : en effet, hormis la collection du département des Estampes, on ne compte qu'une quinzaine d'autres documents de Boullée conservés dans le monde. La numérisation a été effectuée à partir d'ektachromes 24 x 30 cm. Le catalogage a été effectué à la pièce, par numérisation, reconnaissance optique de caractères et



Vue intérieure de la nouvelle salle projetée pour l'agrandissement de la Bibliothèque du Roi

conversion d'un catalogue imprimé en format InterMarc dans BN-Opaline, et les notices ont été regroupées en lots centrés sur un bâtiment ou un projet.

- Dessins d'architecture de Jean-Jacques Lequeu (1757-1825)

713 documents (aquarelle ou lavis) de grand format (jusqu'à 3,5 x 1,5 m). La numérisation a été effectuée à partir d'un microfilm couleur 35 mm. Le catalogage suit exactement les mêmes principes que ceux utilisés pour le projet Boullée.

- Collection Hippolyte Destailleur

4 906 documents (dessin, aquarelle, lavis, gouache, gravure, etc.) de moyen format (40 x 30 cm) représentant des sites de Paris (1 328 documents) ou de province (3 578 documents) du XVI^e au XIX^e siècle. Certains documents étant recto-verso, le total de ce projet atteint plus de 6 000 images. La numérisation a été effectuée à partir de diapositives couleur 24 x 36 cm. Le catalogage a été effectué à la pièce, par numérisation en interne, reconnaissance optique de caractères et conversion en format InterMarc dans BN-Opaline de catalogues imprimés et les notices ont été regroupées par arrondissement pour Paris et par département pour la province.

- département des Manuscrits : quelques documents cartographiques manuscrits.

- Manuscrits médiévaux à peinture

Parmi la sélection de 12 729 pages provenant de 143 manuscrits choisis parmi les plus « représentatifs » des trésors de la Bibliothèque nationale de France, ce « florilège » inclut les documents de la bibliothèque de Charles V, déjà utilisés pour le projet « 1 000 enluminures sur Internet ». La numérisation a été effectuée à partir d'ektachromes 24 x 30 cm ou 4 x 5 pouces pour la majorité des manuscrits et de diapositives couleur 6 x 6 cm.

C'est parmi les manuscrits que se trouve la première application d'imagerie numérique sur l'Internet de la Bibliothèque nationale de France : la page d'accueil du site Web de la BNF contient une base de données graphiques spécifique intitulée « 1 000 enluminures sur Internet ». Cette base de données a été inaugurée en septembre 1995 pour l'inauguration de l'exposition « Culture et pouvoir » à la Bibliothèque du Congrès de Washington. Sept manuscrits à peinture du XIV^e siècle, la période de Charles V, le fondateur de la Bibliothèque nationale, ont été numérisés au taux de 2 000 x 3 000 points par pouce et présentent plus de 850 images zoomables plein écran et accessibles par menus (textes et liens hypertexte vers d'autres textes ou des images, thèmes avec une structure arborescente à 3 niveaux ou feuilletage des manuscrits). Parmi eux, l'« Atlas catalan » par Abraham Cresques, daté de 1375, représente l'état du monde connu à cette époque et constitue le chef-d'œuvre de ce projet. Ce portulan comporte 6 feuilles de vélin, pliées en leur milieu et montées sur des ais de bois, chaque feuille mesurant 64 x 49 cm. L'original a été photographié sur des cibachromes 4 x 5 pouces et numérisé

au taux de 2 000 x 3 000 points par pouce pour les feuilles entières et au taux de 4 000 x 6 000 points par pouce pour les détails. Plus de 50 détails sont disponibles.



Carte de l'Europe, de l'Afrique et du Proche-Orient



Caravane traversant la route de la soie

4. La numérisation des documents cartographiques

Cette communication est centrée sur les collections de documents cartographiques numérisés, parmi lesquelles on trouve des documents parmi les plus intéressants, mais aussi parmi les plus difficiles à numériser et à exploiter correctement, pour trois raisons principales :

- d'une part, les documents les plus intéressants, les plus rares, les moins connus, proviennent souvent des collections spécialisées et ils sont en outre un excellent choix en ce qui concerne la conservation, la communication, la valorisation et les publications,
- d'autre part, les documents cartographiques, en particulier en raison de leur taille et de la précision des informations qu'ils contiennent, soulèvent des problèmes tant pratiques que techniques très difficiles à résoudre pour pouvoir obtenir une bonne qualité et une utilisation efficace des images numériques,
- enfin, à l'inverse d'une page imprimée qui est très « binaire » et ne contient généralement que des points noirs ou blancs, la numérisation de documents

cartographiques relève plus de « l'artisanat », à la fois par la manipulation du document, notamment s'il présente un format qui sort de l'ordinaire, et par l'utilisation de l'image numérique que l'on peut en obtenir.

La Direction des collections, dans laquelle se trouvent les départements qui conservent les collections de la Bibliothèque nationale de France n'est pas seule engagée dans le processus de numérisation : elle partage cette responsabilité avec la Cellule « numérisation » du département de la Bibliothèque numérique et le département de l'Audiovisuel. À l'origine du projet, un des principaux problèmes à résoudre a été de faire admettre aux autres partenaires que les documents cartographiques n'étaient pas tout à fait des documents comme les autres : le marché était prévu pour la numérisation de 300 000 « images » avec un taux de résolution de 2 048 par 3 072 points par pouce. On savait déjà, depuis les tests initiaux, que ce taux était insuffisant pour une bonne utilisation d'images numérisées à partir de documents cartographiques. Ces derniers ont nécessité une procédure particulière afin d'obtenir à l'écran une lecture satisfaisante des toponymes et des détails graphiques : ils ont été numérisés au taux de résolution de 4 000 par 6 000 points par pouce.

Lors de la phase initiale de tests, la question s'est posée de savoir si la numérisation devait être effectuée à partir du document original ou d'un substitut photographique. À l'époque, dans les années 1995-1998, les plus grands numériseurs existant en France ne pouvaient pas traiter de documents à plat de plus de 1,20 m pour les numériseurs à rouleaux ou de plus de 2 m de largeur pour les caméras numériques les plus performantes. Tous les documents sélectionnés dans les projets de numérisation n'atteignent pas de telles dimensions et, dans la plupart des cas, il est possible de numériser un document à partir de l'original, même si les tests montrent que la numérisation d'un inversible couleur de haute qualité donne les mêmes résultats que celle d'un document de taille moyenne. Le problème est différent en ce qui concerne les documents cartographiques dont la taille ne permet pas une numérisation directe : aujourd'hui, il existe des caméras numériques capables de traiter des documents de 4 x 3 m et de restituer l'image en 3D si nécessaire.

En fait, la réponse a plutôt été l'option de la numérisation d'un support photographique : nous étions dans le cadre d'un marché global de numérisation de documents graphiques, parmi lesquels des documents cartographiques et des plans d'architecture. Parmi ces derniers, quelques-uns des plans d'architecture dépassent 3 m de longueur : il est évident qu'ils n'ont pu être numérisés directement et qu'on n'a pas pu faire l'économie d'une campagne photographique. Dans le même ordre d'idées, la carte de Cassini est une mine incomparable en ce qui concerne la

toponymie française du XVIII^e siècle ; il importait donc que l'ensemble des toponymes, quelle que soit la taille de leur typographie, soit distinctement lisible.

La Bibliothèque nationale de France dispose d'une photothèque riche de plusieurs millions de substituts photographiques (microfiches, microfilms, cartes à fenêtre, microfiches monovues, inversibles couleur, inversibles noir et blanc, ektachromes de grand format (jusqu'à 24 x 30 cm), cibachromes, etc.). On aurait pu penser *a priori* que ce trésor allait apporter la solution, mais si l'on en croit les experts en conservation, un inversible couleur réalisé il y a 5 ans n'est plus complètement parfait et un inversible couleur réalisé il y a 20 ans peut être inutilisable, même s'il a été conservé dans un environnement protégé et favorable ; ils soutiennent également que la durée de vie d'une microfiche noir et blanc peut atteindre une centaine d'années, mais il n'était pas question d'utiliser des substituts noir et blanc pour offrir l'image numérique de documents en couleur. On a pu constater que les experts en conservation avaient raison et que nombre de supports photographiques ne présentaient plus un niveau de qualité suffisant. Dans quelques cas, il a été possible d'utiliser des inversibles couleur qui commençaient à virer en leur appliquent une légère correction chromatique selon la charte colorimétrique définie, mais de nombreux documents ont dû faire l'objet d'une nouvelle campagne photographique.

Lors de la sélection des projets, la nature, la couverture et la qualité des reproductions photographiques existantes ont été un important critère de choix : il n'est pas concevable de numériser des cartes à partir de diapositives 35 mm 24 x 36 mm ou de microfiches. Les tests ont montré que les meilleurs résultats étaient obtenus à partir de cibachromes 4 x 5 pouces, d'un piqué plus fin que les ektachromes 24 x 30 cm. Certains projets ont nécessité une campagne photographique complémentaire pour assurer la complétude ou refaire des prises de vues mieux adaptées à une campagne de numérisation que les clichés existants : il a fallu dans certains cas refaire certaines prises de vues de documents dont on pensait la couverture photographique terminée.

Eu égard à la grande quantité d'images à numériser dans le cadre du projet, il est apparu que 300 000 images dépassaient les capacités techniques de la cellule de numérisation et il a été décidé de confier la numérisation proprement dite à des entreprises extérieures dans le cadre d'un marché. Après le lancement d'un appel public à la concurrence et la rédaction d'un cahier des charges, les offres de quatre prestataires ont été retenues, chacun d'entre eux montrant une expertise dans la numérisation de tel ou tel support photographique.

Dans le cas des documents cartographiques, la numérisation a donc été effectuée à partir des cibachromes 4 x 5 pouces au taux de résolution de 4 000 x 6 000 points

par pouce, afin de permettre à l'utilisateur d' « entrer » dans l'image numérique et d'effectuer un zoom sur un détail géographique ou toponymique. Cette procédure, utilisée pour les documents cartographiques, garantit une grande qualité mais coûte nettement plus cher que la numérisation réalisée au taux de 2 000 x 3 000 points par pouce pour les autres documents graphiques.

Dans le cadre particulier des projets de la Bibliothèque nationale de France, le marché passé auprès des 4 prestataires prévoit la numérisation par chacun d'eux de 3 000 images par semaine. En outre, le marché stipule que le prestataire doit numériser en même temps, non seulement l'image elle-même, mais également la « légende ». Celle-ci est généralement composée d'un ensemble de champs extraits de la notice bibliographique, de la référence du cliché au Service de la reproduction qui permet la commande directe d'une reproduction photographique, ainsi que de la clé de la notice bibliographique dans la base de données, utilisée comme lien hypertexte. Le système prévoit, en cas de réponses multiples à une requête, l'organisation des réponses en mosaïques d'images (16 à 32 selon les projets) qu'il est possible, par un simple clic de la souris, d'afficher plein écran et/ou de sélectionner dans un « panier » consultable en fin de requête.

Exemple de légende :

Nolin, Jean-Baptiste (1657-1708). - Roiaume et duché de Septimanie... / dessinée par J.B. Nolin. - [S.l.] : [s.n.], [17e ou 18e siècle]. - 1 carte : 36 x 45,5 cm. - [Cote : BNF C.Pl. Ge DD 2987 (368), microfiche 86/244. - FRANCE-- Sud-Est-- Frontières -- 0561-0752. 16e ou 17e siècle. - SEPTIMANIE. 16e ou 17e siècle. \ Opaline 074046.

Le résultat est tout à fait lisible sur l'écran 17" du poste audiovisuel (PAV), et est encore meilleur sur l'écran graphique 21" du poste d'accès à la bibliothèque numérique (PABN). Malheureusement, le poste audiovisuel est doté d'un imageur qui ne permet pas le zoom au-delà de 200 % et une nouvelle solution technique est à l'étude. Si la consultation d'une image numérique sur un écran graphique ne pose pas problème et permet la lecture des toponymes sur une carte, le problème est la taille gigantesque des fichiers à traiter : une carte ancienne, c'est-à-dire un document pas très détaillé, peut dépasser 400 Mo..., même si la compression permet de le réduire à environ 10 Mo en haute définition ou 300 Ko en basse définition. Certains documents particulièrement détaillés se mesurent en Go...

Cette qualité de consultation a pu être obtenue grâce à un sévère contrôle de qualité tout au long de la chaîne de travail et notamment lors du retour des supports de numérisation : la vérification de la conformité à la charte colorimétrique et la vérification de l'adéquation des légendes à l'image font l'objet d'un examen particulièrement attentif. L'utilisation de normes, à la fois pour la

numérisation, la compression des données, l'exploitation et la communication est également un facteur de qualité. Cette normalisation est un pari sur l'avenir : les données seront toujours plus importantes que les supports physiques. Le support peut évoluer, mais les données restent, notamment si elles possèdent une structure normalisée et sont conservées dans un environnement favorable.

En ce qui concerne le support d'exploitation, c'est-à-dire le serveur, des machines spécifiques, de type Sequant possédant un espace de stockage se chiffrant en téraoctets sont utilisées.

5. Catalogage et indexation

Le problème du traitement bibliographique des documents numérisés a dû être résolu avant les opérations de numérisation. Sur ce point encore, de nombreuses questions se sont posées : l'image numérique est-elle un document à part entière ou un simple support de substitution ? Doit-on cataloguer les images numériques à l'unité ou bien par lots ? Dans le cas des documents cartographiques, la réponse a été de cataloguer tous les documents à la pièce, même s'ils appartiennent à un ensemble plus large, artificiel ou non. En outre, l'image numérique d'un document est considérée comme un substitut de ce document, comme un ektachrome ou un microfilm, ce qui permet de lui faire subir le même traitement intellectuel et la même indexation que le document original.

La solution qui consiste à cataloguer chaque document comme une unité documentaire séparée et à créer des liens entre eux afin d'obtenir un ensemble dans la base de données s'avère longue et coûteuse. Les documents originaux ayant été catalogués à l'unité dans la base de données BN-Opaline, il a suffi d'ajouter une adresse logique aux descriptions bibliographiques et de créer un lien réciproque afin d'obtenir aussi bien l'image numérique à partir de la notice de description bibliographique dans la base de données que la description textuelle à partir de l'image. L'utilisation d'un catalogage capable de gérer des notices analytiques ou un catalogage à niveaux a permis de lier précisément une ou plusieurs images numériques à une notice textuelle contenue dans un ensemble plus important. La conversion rétrospective d'un fichier ou d'un catalogue imprimé, par numérisation et reconnaissance optique de caractères, a été utilisée pour obtenir un catalogage à la pièce. Pour un plus grand confort d'utilisation du système de recherche par l'utilisateur, les projets sont organisés en lots, grâce à des notices de regroupement.

Les ISBD, les normes AFNOR, les formats MARC et les listes raisonnées d'indexation matières (LCSH, RVM ou RAMEAU) ont été utilisés alors que la structuration électronique des données, l'hypertexte, l'indexation en « langage naturel » et l'interrogation en texte intégral sont en plein développement. Les bases

de données disponibles à la Bibliothèque nationale de France ne permettaient pas l'expérimentation de solutions plus modernes. Les projets en cours (« Voyages en France », qui associe textes et images, « Mémoires », qui se veut une vitrine de la diversité des collections de la Bibliothèque nationale de France) sont réalisés grâce à des DTD et une structuration en XML.

La base BN-Opaline a bénéficié dès la fin de 1995 d'un accès Telnet, remplacé en mars 2001 par un accès Web, mais ces accès ne concernent que la base de données elle-même et ne fournissent pas les liens à l'image. C'est pourquoi un appel d'offres a été lancé en mars dernier pour la rédaction d'un schéma directeur d'une nouvelle informatisation des collections spécialisées, dans lequel l'utilisation de métadonnées et le lien à l'image numérique dans une structuration XML seront étudiés.

6. Coût de la numérisation

Bien que le coût de la numérisation soit en baisse constante depuis le milieu des années 1990, le coût total d'une campagne de numérisation est élevé, particulièrement si l'on choisit des taux de résolution importants, mais l'on doit garder à l'esprit que le coût de la numérisation ne comprend pas la seule numérisation, mais couvre également les éventuels coûts de restauration, les éventuelles campagnes photographiques, le catalogage, l'indexation, les différents supports physiques, les postes de travail spécifiques munies d'écrans graphiques « haute définition », etc.

7. Les réseaux

L'établissement a dû définir ses besoins:

- combien d'utilisateurs sont susceptibles de se connecter simultanément ?
Actuellement, 1500 ports sont disponibles simultanément,
- quelles images doit-on rendre disponibles pour l'accès à distance ?
La décision a été prise de ne fournir en accès par l'Internet que les images libres de droits, qui ne représentent pas l'intégralité des fonds numériques disponibles à la consultation sur site,
- combien d'images doit-on rendre disponibles pour l'accès à distance ?
La montée en charge progressive du projet fait que chaque mise à jour apporte son lot de nouveautés à la Bibliothèque numérique,
- quelle qualité d'image numérique doit être disponible sur le réseau ?

Il est tout à fait possible de consulter un document à un taux de définition très élevé à l'intérieur de l'établissement et d'offrir une image numérique de qualité inférieure sur le réseau ; cette question est déterminante et structurante en termes de flux de données et de temps de réponse.

Cette volonté de l'établissement de permettre la fourniture d'images à distance a soulevé des questions : les images sélectionnées sont libres de droits et c'est la Bibliothèque nationale qui détient les droits sur les images numériques. À la fois pour offrir à l'utilisateur des temps de chargement acceptables et pour ne pas risquer le piratage, la Bibliothèque nationale de France a décidé de fournir sur l'Internet des images de qualité dégradée. Si cette procédure est envisageable pour beaucoup de documents graphiques, elle n'est pas intéressante pour les documents cartographiques : en effet, la carte de Cassini qui apparaît en image plein écran non zoomable en qualité dégradée est inutilisable. Ce problème va être réglé par la fourniture de l'image complète en qualité dégradée accompagnée d'extraits en haute définition.

Aujourd'hui, la bibliothèque numérique Gallica fournit aux utilisateurs distants de la Bibliothèque nationale de France près de 50 000 documents graphiques.

8. L'édition de CD-Roms

Par ailleurs, à partir de ses collections de documents numériques, la Bibliothèque nationale de France produit des documents édités sur CD-Roms. La collection «Sources» a pour objectifs de présenter sur CD-Rom une sélection des plus importants documents conservés par la Bibliothèque nationale de France et faisant partie du patrimoine de l'humanité. Ce choix d'édition permet de rendre accessibles des documents qui, du fait de leur état ou de leur fragilité, sont difficilement manipulables et donc rarement consultables.

Chaque document fait l'objet d'une reproduction très fidèle par numérisation à haute résolution et offre les mêmes opportunités de consultation que l'original ; la possibilité d'effectuer des zooms à quatre paliers sur des détails améliore même la consultation que l'on pourrait effectuer sur le document original. Ces «fac-similés électroniques» s'adressent à un large public qui peut ainsi déchiffrer pas à pas le document, puis élargir sa recherche grâce aux qui mettent l'accent sur les thèmes abordés, ainsi que sur le contexte historique et artistique de l'époque à laquelle le document a été réalisé. Ces commentaires sont rédigés en trois ou quatre langues selon les CD-Roms.

Enfin, une dernière partie du CD-Rom propose une série d'outils logiciels qui permet à l'utilisateur de «s'approprier» le document grâce à des «carnets» et à

des « calques ». Les carnets et les calques sont stockés sur le disque dur de l'ordinateur de l'utilisateur et gardent la trace des notes, observations et analyses rédigées par l'utilisateur, ainsi que des extraits d'images ou de textes qui lui semblent particulièrement intéressants à conserver.

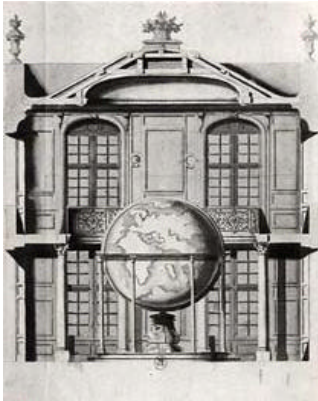
Parmi les documents cartographiques publiés dans la collection « Sources », on peut citer :

- «Mappamondi: une carte du monde au XIVe siècle» qui est une représentation numérique de l'Atlas Catalan de 1375 (Ms Esp 30), avec commentaires en français, anglais, allemand et espagnol, qui reproduit le monde connu à l'époque.



L'atlas catalan (1375)

- «Les globes de Louis XIV: la terre et le ciel par Vincenzo Coronelli», une exploration du globe terrestre et du globe céleste, chacun de quatre mètres de diamètre, réalisés pour Louis XIV à la fin du XVIIe siècle.



Projet de Louis-Étienne Boullée



Détail du globe terrestre



Salon des globes (1782-1900)

«La Géographie d'Idrîsî: un atlas du monde au XIIe siècle» (Manuscrit oriental 2221). En 68 planches et une mappemonde, cet atlas, «divertissement de celui qui est passionné» pour la pérégrination à travers le monde», est replacé dans le contexte du milieu du XIIe siècle et décrit le monde connu en multipliant les itinéraires et les informations aussi bien historiques, économiques que descriptives, voire anecdotiques.

Pierre-Yves Duchemin
National Library of France
Bibliothèque nationale de France
Direction des collections
Mission de coordination
Informatique et numérisation
58 rue de Richelieu
75084 PARIS Cedex 02
FRANCE
Tél. (33 1) 53 79 83 55
pierre-yves.duchemin@bnf.fr.