

La conversion rétrospective à la Bibliothèque Nationale

Jacqueline Solomiac

Conservateur au Centre de coordination bibliographique et technique,
Chef du projet " conversion rétrospective "

Un projet important se met en place depuis peu à la Bibliothèque Nationale : l'informatisation des catalogues antérieurs à 1970. Rien de surprenant à cela : après la vague de l'informatisation, celle de la rétroconversion déferle sur les bibliothèques. Mais il faut reconnaître que l'opération en cours à la Bibliothèque Nationale est particulière par son ampleur et par l'impact qui en est attendu.

I Les prémices

En 1987, la base bibliographique BN-Opale est ouverte au public. Elle contient les notices de toutes les monographies entrées à la BN par le Dépôt légal depuis 1975 et par acquisition ou don depuis 1983, ainsi que celles de tous les périodiques français catalogués depuis 1975. Et déjà une première étape de la conversion rétrospective se dessine : celle du fichier des ouvrages entrés de 1970 à 1979. C'est dans cette continuité qu'apparaît et se développe le thème de la conversion des fichiers plus anciens. Thème sur lequel notre voisin d'outre-Manche offre alors un exemple de dynamisme et de décision.

Le 14 juillet 1988, le Président de la République annonce l'édification d'une Très Grande Bibliothèque destinée à remplacer, non à seconder, la BN. Cet élément décisif va tout accélérer. Dans cette entreprise ambitieuse, la conversion des catalogues de la BN trouve en effet un cadre idéal qui la justifie totale-

ment. On ne peut imaginer la consultation de fiches anciennes dans une bibliothèque d'avant-garde, sinon comme pièces de musée ! Pas plus qu'on ne peut envisager de se passer de ces millions de notices dans le futur Catalogue collectif français. Cette conjoncture exceptionnelle et les moyens financiers qui en découlent - sans oublier la constance avec laquelle la direction de la BN soutient l'entreprise en haut lieu - vont interdire au projet de s'enliser dans les délais et les incertitudes.

Le premier pas dans le concret s'accomplit lorsque l'Administrateur général demande une étude sur le sujet à deux conservateurs du CCBT (Centre de coordination bibliographique et technique), à A.-M. Beaugendre et à moi-même, sous la responsabilité de notre directeur, Marcelle Beaudiquez. La Cellule de conversion rétrospective est née. Après avoir effectué de nombreux sondages dans les catalogues de la BN, après avoir enquêté auprès de ceux qui ont participé à leur constitution et auprès de ceux qui les gèrent actuellement, la Cellule remet le résultat de ses travaux à M. Le Roy Ladurie, en février 1989. Dans ce rapport, plus de 60 fichiers ou catalogues imprimés sont recensés et plus de 6 millions de notices dénombrées ; face à cette masse d'informations, on suggère un traitement en plusieurs phases ; en ce qui concerne le mode de conversion, le recours à des bases existantes est nettement déconseillé, au vu de tests qui ont démontré un taux de recouvrement trop faible dans OCLC comme dans RLIN et l'on

recommande donc la saisie directe des notices ; enfin on ébauche un cahier des charges. Avec cette étude, la notion de conversion rétrospective prend corps à la BN.

I Les débuts du projet

La date-butoir fixée à 1995 pour l'ouverture de la Bibliothèque de France va nous contraindre à une première mise au point du calendrier. Les 6 millions de notices sont réparties sur trois phases :

■ une phase 1 démarrant en 1990

qui comprendrait les quatre plus importants catalogues de la BN, soit un total d'environ 4 millions de notices

- le Catalogue général auteurs personnes physiques (des origines à 1959) en 232 volumes, soit 2 millions de notices ;
- son supplément sur fiches (1024 tiroirs), soit 1 200 000 notices ;
- le catalogue auteurs (de 1960 à 1969) en 23 volumes, soit 550 000 notices ;
- le fichier des périodiques avant 1960 (568 bacs), soit 265 000 notices.

■ une phase 2 démarrant en 1992

qui regrouperait des fichiers quantitativement variés : Anonymes anciens, Collectivités d'avant 1960, Factums, Actes royaux, etc., soit environ plus de 1 500 000 notices.

■ une phase 3 démarrant en 1993

qui regrouperait les notices des ouvrages en caractères non-latins, soit quelque 200 000 unités.

Nous sommes alors à l'été 1989 et il est grand temps d'aborder la Phase 1. Pour assurer la solidité du dossier (vu l'importance des marchés qu'il impliquera), la BN décide de s'adjoindre les services d'une société extérieure. C'est donc avec la collaboration de trois ingénieurs-conseil de chez Silogia que la Cellule élabore le cahier des charges de la phase 1. A l'été 1990, le Secrétariat aux grands travaux ayant donné le feu vert (indispensable car la conversion doit être financée par le budget de l'établissement Bibliothèque de France), l'appel d'offres international peut être lancé. Les étapes de pré-sélection, de bancs d'essai et de sélection finale des soumissionnaires sont franchies tambour battant de l'automne 1990 à janvier 1991, toujours avec l'aide de la société Silogia. Et, deux ans et demi seulement après l'annonce du 14 juillet, les marchés sont signés le 1^{er} mars 1991.

La partition de la phase 1

Pourquoi plusieurs marchés ?

Les principes qui sous-tendent l'appel d'offres sont les suivants :

- ne pas confier, par prudence, la totalité des travaux à une seule société de saisie ;
- envisager, dès le début de l'opération, et sur un plan informatique, la constitution d'une base homogène par rapprochement des notices éparses dans un ou n catalogues (plus exactement le rapprochement des diverses entrées à une même notice) ;
- réduire le plus possible la préparation des fichiers par la Cellule ;
- ne pas laisser à la Cellule seule la charge de gérer et contrôler une opération de cette ampleur et de cette complexité.

A partir de ces axes, six lots contractuels composent l'appel d'offres. Chacun des six CCTP (Cahier des clauses techniques paritaires) va assez loin dans le détail de ce qui est attendu du prestataire :

■ **Les lots 1, 2, 3 et 4** couvrent, respectivement, la saisie du Catalogue général, de son supplément sur fiches, du Catalogue 60-69 et du fichier des périodiques avant 1960.

Chacun de ces cahiers des charges contient :

- la description de la structure du catalogue et de ses notices ;
- les consignes de formatage en INTER-MARC : le prestataire doit identifier les informations contenues dans les notices telles qu'elles se présentent sur le papier, sans réelle préparation de la BN, et les saisir dans les zones INTER-MARC adéquates. Il doit aussi " déduire " certaines informations, comme la langue de l'ouvrage ou le pays d'édition ;
- l'indication du taux de qualité requis, au-delà duquel une tranche de travaux peut être rejetée par la BN, ce qui implique son recyclage complet par le prestataire ;
- la répartition de la charge de saisie sur les 4 années à venir selon des calendriers serrés et rigoureux qui planifient en particulier la circulation des originaux et des produits attendus.

■ **Le lot 5** doit aider à la gestion de l'ensemble et au contrôle de la saisie.

Son cahier des charges contient :

- la définition des outils de gestion qui seront utiles à la BN pour suivre le projet (plannings, statistiques, etc.) ;
- la description des deux types de contrôle qui devront être mis en œuvre chaque mois pour chacun des 4 lots de saisie, afin de valider chaque tranche de travaux effectuée.

Quelques mots sur les deux contrôles éclaireront l'importance de ce lot :

- le premier doit s'assurer de l'exhaustivité de la saisie. Sa performance repose sur le fait que chaque notice saisie est identifiée par un numéro qui la situe au sein de son catalogue d'origine. Ainsi " CG 0003 0127.02 04 " désigne, dans le Catalogue Général, dans le volume 3, à la page 127, sous le 2^e auteur apparaissant sur cette page, la 4^e notice. Cet identifiant, simple à construire, et, nous l'espérons, surtout simple à lire, permettra de remonter à tout moment de l'enregistrement à sa source. Le prestataire du lot 5 a pour tâche de saisir tous les identifiants d'un volume ou d'un tiroir et de les comparer informatiquement aux identifiants fournis par le prestataire de saisie pour le même volume ou tiroir. Seuls les cas litigieux seront tranchés par la Cellule.

- Le deuxième contrôle doit s'assurer de la qualité de la saisie. Le prestataire sélectionne, de façon aléatoire, un échantillon qu'il aura quantifié d'après

la norme AFNOR X06 022. Il en contrôle la frappe, tandis que la cellule en vérifie le formatage. C'est au vu des résultats de ces deux contrôles que la BN décidera chaque mois de la réception ou non d'une tranche de travaux.

■ **Le lot 6**, entièrement constitué de traitements informatiques, doit rassembler et gérer toutes les notices fournies sur bandes par les prestataires des lots 1 à 4, au fur et à mesure de leur saisie, après leur contrôle par le lot 5 et leur validation par la BN. Le regroupement des différents points d'accès d'une notice représente le plus important de ces traitements le critère de la cote servira à rapprocher une notice " secondaire " (dont la vedette est un auteur secondaire) de sa notice " principale " (la même description avec, en vedette, l'auteur principal) et des vérifications supplémentaires sont prévues après rapprochement, sur l'identité des titres essentiellement.

L'enjeu de l'opération n'échappera à personne, pas plus que sa lourdeur, qui ira en augmentant au fil des ans : de sa fiabilité dépendra la cohérence et la " légèreté " de la base.

Un autre aspect non négligeable de ce lot est la production, avec mise à jour semestrielle, de CD-Rom de travail à usage interne, qui offriront à plusieurs services de la BN l'accès aux données de la Conversion rétrospective, ceci dès mi-92, sans attendre la réalisation complète du projet.

Les prestataires de la conversion

■ La saisie

Les lots 1 et 4 ont été attribués aux AIC (Ateliers informatiques du Centre, sis à La Châtre), les lots 2 et 3 aux sociétés co-traitantes Jouve (ateliers à Mayenne) et Saztec (ateliers à Ardrossan, Écosse). L'offre des AIC avait une option non prévue au cahier des charges et qui a été retenue : la numérisation et le stockage sur disque optique numérique des catalogues originaux.

■ Le contrôle

Le lot 5 a été confié au Bureau Van Dijk (Paris), qui proposait un contrôle qualité plus complet que celui de notre cahier des charges : l'échantillon sélectionné

serait à nouveau saisi par une société sous-traitante (Cégitex) et une comparaison informatique serait faite entre cette saisie et les notices produites par Jouve, Saztec et les AIC. A la Cellule resteraient le soin d'examiner les différences révélées par la confrontation et la responsabilité d'accepter ou de rejeter une tranche de travaux.

Le lot 6, la constitution finale de la base, a été confié à la société Jouve.

Où en est la conversion à l'automne 1991 ?

Tous les lots ne démarrent pas ensemble, bien sûr, mais tous suivent un planning similaire : d'abord quelques mois de mise au point des spécifications et de concertation avec le prestataire (deux mois pour un lot de saisie, six pour le lot 6), puis une période de tests, enfin, lorsque les essais sont positifs, la production commence véritablement. La montée en charge s'échelonne de mars à novembre 1991 et, à l'heure où paraîtra cet article, la totalité de la phase 1 sera enclenchée. Aujourd'hui, c'est plus de 150 000 notices du Catalogue général qui ont été saisies, dont un bon nombre ont été contrôlées par le lot 5 ; le lot 2 est en période de tests ; la dernière touche (sur le papier au moins) est donnée aux spécifications du lot 6 et les premières réunions de mise en place des lots 3 et 4 vont se tenir.

Jusqu'à présent donc, les calendriers, tout draconiens qu'ils sont, ont été respectés ! Mais la route est longue jusqu'en 1995 et aucune aire de repos n'est prévue !

Les tâches de la cellule

■ Les écueils de la conversion

Pour répondre à la charge de travail croissante, l'effectif de la Cellule est passé de 2 personnes en 1988 à 13 en août 1991 (4 conservateurs, 3 bibliothécaires-adjoints, 6 contractuels engagés sur le budget Bibliothèque de France, 3 de catégorie A, 3 de catégorie B), sans compter 2 agents de catégorie C à mi-temps, employés à composer les fiches du lot 2. Marcelle Beaudiquez,

quant à elle, dirige l'ensemble du projet et en assume la part administrative.

L'ampleur de l'opération entraîne, au quotidien, nombre de problèmes d'organisation : la mise en place et le suivi des six lots, les relations avec quatre prestataires différents, la maintenance de la coordination entre les lots, la gestion matérielle des catalogues originaux, le déroulement des recettes successives, toutes ces tâches consomment une bonne part du temps et de l'énergie de la Cellule, d'autant que les délais ne laissent aucun répit ! Rares sont les jours, jusqu'en 1995, qui ne sont pas des dates "contractuelles" !

Les autres difficultés rencontrées, et ce ne sont pas les moindres, tiennent à l'objet même de la conversion, c'est-à-dire aux catalogues : le Catalogue général en 232 volumes et son Supplément sur fiches en sont particulièrement fertiles, et de façon bien différente : le Catalogue général, lui, s'est voulu dès l'origine cohérent, raisonné, "normalisé" et il l'est ! C'est-à-dire qu'il est à la fois une mine bibliographique et un modèle de concision et d'économie : par exemple, il offre des séquences de notices dont seule la première est complète, les suivantes, autres éditions de la même œuvre, ne comportant que les éléments divergents. Dans ces cas, pour que chaque notice ait une vie autonome au sein d'une base, il faut recréer, par duplication, l'information sous-entendue et donc défaire ce que les catalogueurs du passé avaient si savamment construit. Quand des dépouillements de suites à plusieurs niveaux sont agrémentés de "blocages" de cette sorte, le retour aux informations originales devient un sérieux casse-tête. De plus, tout soigneusement préparé et imprimé que ce catalogue ait été, il abonde en corrections et ajouts manuscrits, pas toujours déchiffrables hélas ! Quant au Supplément, c'est la variété qui le caractérise. Il présente un panorama impressionnant de plus d'un siècle de catalogage, et...de catalogueurs ! Et sans un seul coup de peigne d'ensemble ! Les notices y sont disparates sur le plan matériel : manuscrites, imprimées, dactylographiées, petites cartes photocopiées, bribes de bibliographies découpées et collées, notices amputées de leurs débuts de lignes par un massicotage malheureux, etc. Et très diverses dans leur structure : notices

réduites à un auteur plus une cote, notices plus que complètes, avec biographie de l'auteur ou indexations obsolètes, notices sans cote... Tout se rencontre ! Et les noms des auteurs connaissent les aléas de la multiplicité de formes, alors que, sur ce plan, le Catalogue général demeure encore souvent une référence pour les Autorités personnes physiques, avec ses vedettes uniformisées et ses homonymes différenciés.

Dans ce contexte, il était impossible de laisser les prestataires de saisie aborder seuls les catalogues, comme il était impensable de rédiger des consignes pour tous les cas de figure. La Cellule intervient donc de deux façons :

- avant le départ à la saisie, en survolant les volumes ou les tiroirs pour repérer et résoudre les difficultés (notices illisibles, cotes absentes, formes d'auteurs variées, blocages de dépouillements...);
- pendant la saisie, en répondant aux interrogations et aux doutes des prestataires, selon un circuit d'anomalies rigoureusement planifié.

Cependant ces procédures ne peuvent espérer pallier la totalité des aberrations et des insuffisances : il faut donc, dès aujourd'hui, accepter l'idée que la Base de la conversion ne sera ni totalement harmonieuse ni parfaitement cohérente. Par contre, les défauts et les failles de cette base seront largement compensés par ce qu'à l'évidence, elle offrira : des accès multipliés et facilités à une formidable masse d'informations.

Conclusion ?

La véritable conclusion s'écrira en 1995. C'est à ce moment que tous les chantiers Bibliothèque de France convergeront pour aboutir à leur objectif : la mise à la disposition du plus grand nombre de notre patrimoine éditorial. C'est également à cette date que notre travail devrait entrer dans sa dernière phase, l'exploitation et la diffusion de la Base de la conversion rétrospective, et y trouver sa pleine justification. Car nous savons combien ces informations bibliographiques sont attendues de tous, lecteurs et bibliothécaires.