

L'opération n.c.g. ou l'intégration du catalogue général 1970-1979 dans la base bibliographique bn-opale

*par Annick BERNARD, Conservateur, Chef de
projet et Monique CHOUDEY, Conservateur, Chef
du Service du Nouveau Catalogue Général du
Département des Livres imprimés.*

Le Nouveau Catalogue Général

à la fin des années 1950, la B.N. décidait d'arrêter à 1959 le contenu des volumes du Catalogue général des livres imprimés (1) restant à paraître, et de publier dans une nouvelle série les notices des ouvrages entrés à partir de 1960. A cette fin, le service du Nouveau Catalogue Général était créé en février 1960.

La première tranche quinquennale de la nouvelle série, recensant les ouvrages entrés de 1960 à 1964, était publiée de 1965 à 1967.

Cette tranche était ensuite reprise dans le catalogue décennal 1960-1969 publié de 1972 à 1978.

En raison de l'accroissement considérable de la masse documentaire au cours de la décennie 1970, le fichier constitué des notices des ouvrages entrés de 1970 à 1979 atteignait 1 million de fiches.

Si la série en caractères latins du catalogue 1960-1969 était complète en 23 volumes obtenus par photographie des fiches, la reproduction de la tranche 1970-1979 par le même procédé aurait produit un catalogue d'une cinquantaine de volumes, d'un prix de revient évalué à près de 10 millions de francs. C'était une somme énorme, et cependant, au début des années 1980, il devenait urgent de prendre une décision que le service du Nouveau Catalogue Général attendait pour organiser son travail.

Or, au moment où la Bibliothèque nationale établissait son schéma directeur in-

formatique qui devait aboutir en juillet 1982 à la décision de créer une grande base bibliographique, une société de composition proposait une alternative au procédé traditionnel : celle-ci consistait à enregistrer les notices qui pourraient ensuite être reversées dans la base.

L'installation prévue par cette société d'un atelier de saisie en Afrique du Nord devant permettre de réduire les coûts de moitié par rapport à ceux du catalogue obtenu par reproduction des fiches.

L'idée fut retenue dans son principe. Cependant, chacun était conscient qu'une étude sérieuse était nécessaire avant de lancer l'opération.

Le problème n'était pas simple. Le fichier n'avait pas été constitué pour l'objectif qu'on voulait lui assigner maintenant : en effet, les différentes entrées d'une même notice se trouvaient disséminées dans l'ordre alphabétique comme elles doivent l'être dans un catalogue par noms d'auteurs. Or, pour pouvoir alimenter la base bibliographique, il fallait réaliser l'opération inverse et rattacher les entrées secondaires à la notice principale. En l'absence d'un numéro B.N. unique du type LC number (2), la cote des ouvrages était le seul moyen fiable pour procéder à cette réunification.

De plus, l'édition de la *Bibliographie de la France* étant automatisée depuis 1975, il aurait été absurde de saisir les fiches correspondant aux notices déjà enregistrées. Celles-ci devaient être reconstituées par programme à partir des

(1) Catalogue général des livres imprimés de la Bibliothèque nationale. Auteurs. - Paris : [Bibliothèque nationale], 1897-1981. - 213 tomes,

232 volumes.

(2) Numéro d'identification de la notice à la Library of Congress

enregistrements en format INTER-MARC.

Enfin, dernière contrainte, le fichier ayant été préalablement classé dans l'ordre alphabétique en vue de l'édition traditionnelle, il importait de sauvegarder ce classement.

Cependant, compte tenu du coût de l'édition traditionnelle, de la création du Service informatique, et des progrès de l'établissement dans la voie de l'automatisation, ainsi que des contrats pris par le N.C.G. avec un imprimeur utilisant lui-même une technique évoluée, il devenait certain que la solution résidait dans le recours à l'informatique, qui permettrait de rentabiliser l'opération en la faisant servir à deux fins.

Le Cahier des charges :

Réalisé au cours du 2ème trimestre 1984 par le Service Informatique et Organisation en étroite collaboration avec le service du N.C.G., le cahier des charges répartissait de la façon suivante les prestations attendues :

lot 1 : a) constitution d'un fichier d'archivage sur support magnétique contenant en format INTERMARC les notices n'existant pas encore sous forme enregistrée, soit :

300 à 350 000 notices principales : ouvrages annoncés dans la Bibliographie de la France de 1970 à 1974 ; ouvrages étrangers entrés de 1970 à 1979 ; publications officielles, acquisitions des départements spécialisés ; dons.

350 à 400 000 entrées secondaires.

100 000 fiches de renvoi.

Les 200 000 fiches Canac (3) dont il ne devait être saisi qu'une clé permettant d'en reprendre les éléments dans le fichier magnétique existant n'étaient l'objet d'aucun archivage.

b) fourniture d'un fichier réunissant dans un format d'édition les 800.000 notices saisies plus les 200.000 notices reprises du fichier existant et conservant le tout dans l'ordre alphabétique du fichier d'origine.

lot 2 : production du catalogue :

compte tenu des vœux et des ressources budgétaires des divers utilisateurs du

catalogue, plusieurs hypothèses devaient être envisagées :

- édition sur papier (par photocomposition ou impression laser) ;

- édition sur microfiches (traditionnelles ou COM) ;

- édition mixte combinant un petit tirage sur papier et le reste sur microfiches.

Pour l'enregistrement des notices, les sociétés étaient invitées à rechercher une solution permettant d'alléger au maximum la charge de travail à fournir par le N.C.G. sur le plan de la préparation matérielle des fiches et du contrôle de la saisie. Il leur était suggéré d'examiner l'intérêt d'un recours à la lecture optique. L'appel d'offres :

Comme il convenait pour une opération de cette importance, un appel d'offres ouvert était annoncé en juillet 1984 dans le *Bulletin officiel des annonces des marchés publics*, ainsi que dans le *journal officiel des Communautés européennes*. Les sociétés intéressées devaient fournir une réponse pour un lot seulement ou pour l'ensemble des lots.

Les réponses :

Sur les 31 sociétés qui s'étaient présentées pour retirer le cahier des charges, 7 seulement remirent une proposition. Les réponses se répartissaient comme suit :

- Lots 1 et 2 (saisie et impression) : deux réponses ;

- Lot 1 seul (saisie) : deux réponses ;

- Lot 2 seul (impression) : trois réponses.

Il apparut très vite que sur les sept réponses, nous pouvions retenir, dans une première sélection, quatre propositions sérieusement élaborées qui méritaient d'être analysées finement. Elles provenaient :

- de deux sociétés qui répondaient pour l'ensemble des lots (sociétés A et B),

- de l'une des sociétés de saisie répondant pour le lot 1 seul (société C),

- d'un imprimeur répondant pour le lot 2 (société D).

Des réunions de travail furent organisées avec des représentants de ces différentes sociétés pour obtenir des précisions sur des points obscurs ou douteux des réponses, et aussi pour s'assurer que nos spécifications avaient été bien com-

prises.

La société A proposait pour le lot 1 deux options, l'une classique de saisie puis relecture, l'autre plus originale qui consistait à tenter d'obtenir, au moyen d'une clé de recherche qui serait saisie, un grand nombre de notices existant déjà dans une base commercialisée par un distributeur spécialisé¹.

Quant à l'usage de la lecture optique, la diversité de présentation et de typographie des notices d'origines variées qui constituaient le fichier conduisit les sociétés qui l'avaient envisagé à y renoncer (4).

Les Bancs d'essai :

A la fin de novembre 1984, la présélection fut officiellement annoncée, et les trois sociétés retenues pour le lot 1 avisées qu'elles devaient se soumettre au banc d'essai annoncé dans le cahier des charges.

A cette fin, un lot de 200 notices représentatives de l'ensemble du fichier y compris pour le nombre et la variété des caractères accentués - fut préparé par le service du N.C.G. et, pour les sociétés B et C, divisé en deux sous-lots de 100 notices. Le premier sous-lot fut remis aux sociétés une semaine avant le banc d'essai pour leur permettre d'évaluer la charge de travail, d'établir le format de saisie nécessaire, et d'entraîner leur personnel.

Le deuxième sous-lot, apporté par les représentants de la BN le jour même du banc d'essai, devait être saisi en double frappe en deux heures et les états de contrôle, comportant les signes diacritiques saisis, remis à la B.N. le lendemain de l'essai au plus tard.

Le Choix :

Les bancs d'essai furent très instructifs. Si les résultats de l'essai de saisie des sociétés B et C se montrèrent d'une qualité équivalente qui n'aurait pas permis à elle seule de se déterminer, le fait de se rendre sur les lieux fit apparaître le principal point faible de la proposition de la société B, dont le sous-traitant ne disposait en fait que d'une entreprise de petite taille, certes très performante et utilisée au mieux de ses possibilités, mais à laquelle on pouvait hésiter à confier un

(3) Canac (Catalogue national centralisé) : service de distribution de fiches produites sur imprimante à partir de la Bibliographie de la France automatisée. A fonctionné de 1975 à 1984. Une solution transitoire a été mise en place en attendant que les fiches puissent être produites à partir de la base BN-Opale.
(4) Le cahier des charges contenait en annexe la

photographie d'un choix d'exemples représentatifs. Rappelons pour expliquer cette diversité que :
- un changement de norme est intervenu en 1975 à la suite de la publication de l'I.S.B.D.
- le catalogue de la BN rassemble des fiches imprimées ou multigraphiées provenant des divers circuits de catalogage (français, étranger, langues

rares, départements spécialisés).

(5) Dues sans doute à la provenance d'une partie des notices.

(6) L'utilisation du fichier tel qu'il était constitué, sans modification du classement. Ces différences auraient pu être acceptées - après qu'on en ait sérieusement pesé les conséquences - s'il s'était agi

travail d'un volume important nécessitant un rythme soutenu et régulier comme c'était le cas du catalogue 1970-1979.

La société C au contraire confirma la bonne impression que l'on avait de sa solidité et de son organisation. Elle disposait en province d'un atelier de 70 personnes et de liaisons très efficaces avec Paris.

Quant à la société A, la recherche dans la base en format MARC produisit des résultats intéressants puisque 50 % des notices furent trouvées (au lieu de 60 % annoncés) et que l'écart de catalogage entre ces notices et les nôtres n'était pas considérable. Toutefois, des différences dans les vedettes, des lacunes (5) aussi importantes que la mention de collection dans un certain nombre de notices et enfin la forme anglaise des notes, rendaient cette solution impraticable compte tenu du principe de départ adopté pour la réalisation du catalogue 1970-1979 (6) et des contrôles et modifications nécessaires pour s'y conformer (7).

De plus, la reprise des notices d'origine étrangère, admissible pour les ouvrages entrés par achat, don ou échange, était contestable pour les notices d'ouvrages entrés par dépôt légal (8).

Enfin, la proposition ne présentait pas d'avantage sur le plan financier par rapport aux deux autres propositions comparables.

La société A.I.C. (Ateliers informatique du Centre), appelée, ci-dessus société C, dont la proposition se montrait la plus sûre sur tous les plans (9) fut donc choisie en février 1985 et un marché portant sur trois exercices conclu avec elle quelque temps plus tard.

La B.N. s'engageait à remettre après une période de montée en charge, 50 tiroirs par mois jusqu'à épuisement du fichier, la saisie devant se terminer à l'automne 1987.

La Réalisation :

La première étape consistait à analyser finement le fichier pour établir une typologie dont dériveraient les consignes de saisie et de marquage des zones par le personnel des A.I.C. Les notices selon leur type furent réparties en catégories

affectées d'un statut (10). Les différentes présentations des zones vedettes et de la zone titre, qui varient selon l'origine des notices et la norme en vigueur à l'époque de leur rédaction, furent recensées et l'ensemble ramené à un certain nombre de cas entraînant l'application de règles précises. A la charge du service du N.C.G. restait le repérage des cas douteux ou atypiques.

Le marquage des fiches constitue une première étape dans la structuration des notices en format INTERMARC. Les éléments reconnus à ce stade sont : les vedettes auteurs, collectivités- auteurs et vedettes titres, le titre propre, la mention de collection dans les notices conformes à l'ISBD, les notes en bloc, l'ISBN, le numéro d'entrée, la cote des différents exemplaires, les entrées secondaires selon leur nature, la tomaison dans le cas des dépouillements. La partie du corps de la notice qui suit le titre propre, des compléments du titre à la collation incluse, n'est pas structurée.

La Chaîne de traitement :

Un chaîne de traitement en plusieurs modules a été réalisée par la société AIC pour effectuer par programme :

- le rapprochement des entrées principales et secondaires des notices au moyen de la cote ou d'une clé en tenant lieu dans certains cas,
- le rapprochement des entrées principales et secondaires CANAC ; la reprise des notices 1975-1979 de la *Bibliographie de la France* dans le fichier INTERMARC correspondant,
- la mise à jour (en différé) de la saisie : il est possible à toute époque de modifier les notices déjà enregistrées par les AIC et de changer, s'il y a lieu, leur place dans l'ordre alphabétique,
- la conversion en format INTERMARC. Les traitements sont effectués au centre informatique de la B.N., sous la responsabilité des A.I.C.

Le travail au service du N.C.G.

La solution adoptée n'a pas été - on s'en doute - sans conséquences sur l'organisation du travail au sein du service.

techniques qui avaient été pris par les sociétés A et B rendaient difficiles d'évaluer à l'avance certains coûts.

(10) Principale, Secondaire, Renvoi.

Notice simple, "chapeau" ou notice de premier niveau, dépouillement, Canac.

(11) Précisons qu'il s'agit d'une édition au format

Si le personnel effectue des travaux bibliographiques indépendants du mode de production du catalogue : vérification du classement, recherches sur les auteurs homonymes, élimination des doubles traitements, harmonisation des traitements notamment dans le cas des notices à niveaux, il doit aussi faire sa part de la préparation de la saisie en s'assurant que le statut et les éléments de la notice sont parfaitement identifiables par le personnel des A.I.C.

Dans le cas contraire, les codes ou les repères nécessaires sont indiqués par le N.C.G.

Lui incombent aussi la révision des états de contrôle qui accompagnent les tiroirs au retour de la saisie, et celle des comptes rendus des traitements informatiques.

Pour assurer ces différentes tâches, le personnel du N.C.G. a été réparti en équipes :

- deux équipes (9 personnes au total) pour la révision des tiroirs (chaque tiroir est lu deux fois par deux personnes différentes),
- une équipe (4 personnes) chargée du contrôle de l'image de la saisie (11) et des états informatiques issus du rapprochement des notices principales et secondaires.

L'envoi des 50 tiroirs mensuels est généralement réparti en quatre fois, ce qui présente l'avantage de ne pas priver trop longtemps le N.C.G. de tiroirs dont on peut avoir besoin pour effectuer des vérifications.

A l'issue du traitement trimestriel de rapprochement des notices principales et secondaires, le service reçoit plusieurs types d'états informatiques :

- un état des notices apparemment bien rapprochées dont la vérification est assez rapide (on trouve quelques "mauvais rapprochements"),
- des états d'anomalies (doublons de cote, incompatibilité de statut entre notice principale et notice secondaire, etc.).

De même, les traitements de reprise des notices de la *Bibliographie de la France* dans le fichier INTERMARC génèrent des états similaires de contrôle et d'anomalies.

Les anomalies sont traitées par le person-

A4 par imprimante à laser bien plus agréable et commode à manipuler que les sorties d'imprimante classique. Après relecture, les corrections sont renvoyées aux A.I.C. et effectuées avant l'intégration des lots de saisie au fichier de traitement.

uniquement d'enrichir la base de données aux moindres frais.

(7) Sans oublier la conversion nécessaire du format MARC à INTERMARC.

(8) Ce catalogue sert de cumulatif de la *Bibliographie de la France*.

(9) Y compris sur le plan financier car les partis

nel des deux équipes de révision du fichier. Les erreurs détectées par programme, comme celles qui ont pu être repérées au cours du travail par le programme de mise à jour avant le traitement de rapprochement suivant.

La Conversion :

Les notices sont, nous l'avons vu, saisies en format INTERMARC rudimentaire. Des spécifications très précises, tenant compte de l'expérience déjà acquise par la B.N. dans ce domaine, ont été établies. Par programme, certains codages sont affectés aux notices ; le découpage des zones en sous-zones découle soit de la ponctuation, soit des repères saisis ou encore de l'analyse de certaines chaînes de caractères (12).

Trois fichiers différents résultent de cette conversion : un fichier INTERMARC (M) pour les monographies, un fichier INTERMARC (S) pour les publications en série (13), et un fichier INTERMARC (A), c'est-à-dire en format de la notice d'autorité, créé à partir des notices de renvoi de vedette à vedette.

Le Chargement dans la base :

Dans le courant de 1988, les lots de notices disponibles seront intégrés à la base BN-Opale selon la procédure normale : indexation des zones qui constituent des points d'accès, une comparaison des zones vedettes par rapport aux fichiers d'autorité, création de notices d'autorité "provisaires" (14), chaînage des notices de premier et de deuxième niveau (chapeaux/dépouillements), chaînage des notices bibliographiques aux notices d'autorité.

Les notices seront marquées d'un statut signifiant qu'elles ont - du point de vue de la base - un degré d'analyse inférieur à celui de notices créées directement en format INTERMARC. En effet, si les points d'accès essentiels sont préservés : vedettes, titres-propre, titre de collection (15), ISBN, ISSN, numéro d'entrée, cote, la structure de la notice est moins fine et l'affectation automatique de certains codes relativement imprécise.

Grâce au numéro de saisie servant de lien, les notices déjà chargées pourront être enrichies des entrées secondaires

rencontrées dans la partie du fichier 1970-1979 traitée ultérieurement.

Après les apurements nécessaires qui seront effectués après la fin de la saisie : réduction des anomalies subsistantes, traitement des notices secondaires qui n'auraient pas trouvé leur mère, le chargement sera vraisemblablement terminé au printemps 1988.

L'édition du Catalogue :

Selon toute probabilité, conformément au vœu de la majorité des acheteurs potentiels consultés par questionnaire, le catalogue sera publié sur microfiches COM, procédé économique et, de surcroît, cohérent avec la décision prise par la BN de faire reproduire sur microfiches le Catalogue général imprimé, son supplément sur fiches, et le Catalogue 1960-1969 (16).

Une édition mixte - partie sur papier, partie sur microfiches - envisagée un moment car plus satisfaisante du point de vue des utilisateurs, est, à ce jour, à exclure, car trop onéreuse.

Conclusion - Pourquoi cela marche :

On peut dire après coup que c'était une gageure de lancer cette opération coûteuse à un moment où nous ne pouvions pas savoir si toutes les conditions étaient réunies pour son succès.

Les difficultés ont été surmontées les unes après les autres et l'on peut se féliciter pour différentes raisons dont nous donnerons quelques-unes :

- d'avoir trouvé dans la société A.I.C. des interlocuteurs qui ont parfaitement compris le problème posé et ont su constamment s'adapter aux besoins de la B.N.

- d'avoir trouvé dans le service du N.C.G. une volonté de respecter les engagements de l'établissement et une motivation - soutenue il est vrai par des résultats tangibles - qui persiste au fil des mois malgré un travail souvent ingrat.

La stabilité des personnes dans les trois secteurs concernés - service rédacteur du catalogue, service informatique et société de service - est aussi un facteur non négligeable de succès. Elle a permis de maintenir la collaboration fructueuse

qui s'est instaurée dès le début de l'opération entre les trois partenaires.

Du côté financier, la B.N. a su faire l'effort nécessaire ; pour un coût moindre que celui de la simple reproduction photographique du fichier, l'établissement aura réalisé l'édition du catalogue et presque doublé le volume initial de la base bibliographique.

La cohérence future de celle-ci - nous voulons dire au moment où le catalogue sera chargé - est préservée au mieux des possibilités par les relations étroites et constantes du service du N.C.G. avec les responsables des fichiers d'autorité noms de personne et collectivités (17).

Ainsi grâce à une décision prise en 1982 avec les yeux de la foi et le secours de la logique, la B.N. disposera en 1989 d'une base complète pour les livres français et étrangers entrés depuis 1970 (18).

L'ensemble sera accessible aux lecteurs, qui bénéficieront d'une richesse d'accès que ne permettent pas les catalogues manuels. Les catalogueurs auront à leur portée immédiate un instrument qui facilitera considérablement leur travail. Enfin, les bibliothèques pourront obtenir les notices pour leur propre catalogue grâce à plusieurs clés d'accès (19) et, dans un futur proche, les consulter en ligne.

Perspectives d'avenir :

Selon les moyens dont elle disposera et les possibilités offertes par les nouvelles technologies, plusieurs voies s'offriront à la Bibliothèque nationale :

- à l'instar de la British Library (20), faire enregistrer intégralement son catalogue imprimé,

- s'en tenir dans un premier temps à un enrichissement sélectif par la mise à profit de toutes les occasions : saisie de catalogues destinés à la publication, apport de départements spécialisés (dont les livres ne sont pas intégralement dans le N.C.G.), traitement des livres anciens (21), apport du Centre de prêt.

(13) De nombreuses publications officielles sont en fait des publications en série.

(14) C'est-à-dire créées automatiquement. Elles garderont ce statut jusqu'à leur validation éventuelle.

(15) Sauf pour les notices 1970-1974 antérieures à l'application de l'ISBD ; la mention de collection se trouvant au début des notes, il n'est pas possible de l'isoler à coup sûr au moment de la préparation de la saisie par du personnel non bibliothécaire.

(16) Les trois catalogues ont été publiés par Chawck-Healey France en 1987.

(17) Relations rendues plus faciles par l'emménagement du NCG dans le bâtiment Vivienne.

(18) Sous réserve de quelques compléments de saisie déjà prévus, notamment pour les livres étrangers entrés pendant la période 1980-1983.

(19) Clés numériques permettant d'obtenir les notices : ISBN, numéro d'annonce dans la *Bibliogra-*

phie de la France (à partir de 1975) ou numéro B.N. (numéro attribué par le système de gestion de la base bibliographique).

(20) Qui vient de décider la saisie des six millions de notices de son catalogue imprimé.

(21) Déjà en cours pour les anonymes anciens enregistrés en format INTERMARC.

(12) Ainsi la zone auteur est découpée : en nom, prénom, qualificatif de date, qualificatif de fonction code de fonction.