



HAL
open science

**Comment assurer la cohérence et la qualité de notices
rétroconverties ? Identification des erreurs et corrections
des notices issues de la conversion rétrospective à la
Bibliothèque nationale de France des documents
imprimés dans BN-Opale Plus**

Olivier Jacquot

► **To cite this version:**

Olivier Jacquot. Comment assurer la cohérence et la qualité de notices rétroconverties ? Identification des erreurs et corrections des notices issues de la conversion rétrospective à la Bibliothèque nationale de France des documents imprimés dans BN-Opale Plus. 1999, 1 vol. (55 p.). hal-04366085

HAL Id: hal-04366085

<https://bnf.hal.science/hal-04366085>

Submitted on 28 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ENSSIB - Grandclément

Olivier JACQUOT

Comment assurer la cohérence et la qualité de notices
rétroconverties ?

Identification des erreurs et corrections des notices
issues de la conversion rétrospective
à la Bibliothèque nationale de France
des documents imprimés dans BN-Opale Plus

Projet Professionnel Personnel de bibliothécaire :
gestion de projet

1999

Sous la direction de Madame Suzanne Peters,
tuteur pédagogique

Conseiller professionnel
d'établissement
Irène Paillard
Beaugendre
Conservateur en chef
BDIC

Correspondant

Anne-Marie

Conservateur général

BnF

Table des matières

| | |
|---|-------------|
| INTRODUCTION | 4 |
| 1. CADRE FONCTIONNEL DU PROJET | 5 |
| 1.1. SITUATIONS HIERARCHIQUES AU SEIN DE LA BNF..... | 5 |
| • <i>La Direction des Services et des Réseaux (DSR)</i> | 5 |
| • <i>L'Agence Bibliographique Nationale (ABN)</i> | 6 |
| • <i>Le Service de la Coordination Bibliographique (SCB, ex-CCBT)</i> | 6 |
| • <i>Le Bureau Qualité du Catalogue (BQC)</i> | 7 |
| 1.2. LE PROJET DANS LES OBJECTIFS GENERAUX DE LA BNF..... | 7 |
| 2. MISE EN PLACE D'UNE POLITIQUE DE COHERENCE DES DONNEES RETROCONVERTIES .. | 7 |
| 2.1. LES RAISONS DU CONTROLE QUALITE DES NOTICES..... | 7 |
| 2.1.1. | <u>Une</u> |
| <u>exigence de qualité</u> | 8 |
| 2.1.2. | |
| <u>L'amélioration des résultats de la recherche</u> | 8 |
| 2.1.3. | |
| <u>L'uniformisation des données qui servent de référence aux bibliothèques</u> . | 9 |
| 2.1.4. | |
| <u>L'obtention de produits d'édition corrects</u> | 9 |
| 2.1.5. | <u>La</u> |
| <u>préservation de l'avenir</u> | 9 |
| 2.2. LES CONTROLES DE LA QUALITE DEJA EFFECTUES SUR LES DONNEES DU CATALOGUE | |
| 10 | |
| 2.2.1. | <u>Lors</u> |
| <u>du catalogage en ligne</u> | 10 |
| 2.2.2. | <u>Lors</u> |
| <u>de la conversion rétrospective</u> | 10 |
| 2.3. LE CONTROLE DE LA QUALITE A MENER SUR LE CATALOGUE BN-OPALE PLUS.... | 12 |
| 3. L'IDENTIFICATION DES PROBLEMES | 13 |
| 3.1. LES ERREURS A CORRIGER DEJA IDENTIFIEES ET LEUR VOLUME..... | 14 |
| 3.1.1. | <u>Les</u> |
| <u>erreurs de la rétroconversion</u> | 14 |
| • <i>Des erreurs originales</i> | 14 |
| • <i>Des erreurs de saisie ou de code</i> | 15 |
| • <i>Des doublons</i> | 16 |
| • <i>Des notices à erreurs multiples</i> | 17 |
| 3.1.2. | <u>Les</u> |
| <u>erreurs de BN-Opale, Sycomore ou de la migration</u> | 17 |
| • <i>Dans les données provenant de BN-Opale</i> | 17 |
| • <i>Du récolement</i> | 17 |
| • <i>De la migration</i> | 18 |
| 3.2. LES PROBLEMES INCONNUS A IDENTIFIER ET COMMENT..... | 18 |
| 3.3. LA VOLUMETRIE DES ERREURS..... | 19 |
| 3.3.1. | <u>De</u> |
| <u>la conversion</u> | 19 |
| • <i>Les erreurs de cotes</i> | 20 |

| | |
|---|------|
| • <i>Les doublons</i> | 21 |
| • <i>Les autres erreurs</i> | 21 |
| 3.3.2..... | Hors |
| <u>conversion</u> | 21 |
| 4. L'AMELIORATION DU CATALOGUE | 22 |
| 4.1. LA CORRECTION DES ERREURS : METHODE ET ORGANISATION DU TRAVAIL..... | 22 |
| 4.1.1..... | Les |
| <u>corrections informatiques</u> | 22 |
| 4.1.2..... | Les |
| <u>corrections mixtes</u> | 22 |
| 4.1.3..... | Les |
| <u>corrections manuelles</u> | 23 |
| 4.2. LES TRAVAUX SUPPLEMENTAIRES..... | 24 |
| 4.2.1..... | Des |
| <u>compléments de conversion</u> | 24 |
| 4.2.2..... | Le |
| <u>catalogage des UC</u> | 25 |
| 4.2.3. | |
| <u>L'enrichissement des notices de la base</u> | 26 |
| 4.3. PROCEDURES DE TRAVAIL..... | 26 |
| 5. LES AGENTS DU BUREAU ET LEURS ATTRIBUTIONS | 28 |
| 5.1. LES COMPETENCES REQUISES..... | 28 |
| 5.2. L'IMPLICATION EN PERSONNEL..... | 29 |
| 5.3. LES FORMATIONS A METTRE EN ŒUVRE..... | 29 |
| 5.4. LES RESPONSABLES DES TRAVAUX DE CORRECTION..... | 30 |
| 5.5. LES COLLABORATEURS DU BQC..... | 31 |
| 5.6. L'ECHEANCIER DES TRAVAUX..... | 32 |
| 6. DES REFERENCES EXTERIEURES : EXEMPLES DE LA BDIC ET DE la BPI | 34 |
| 6.1. ANALYSE DE CES REFERENCES..... | 34 |
| 6.1.1..... | Les |
| <u>convergences</u> | 34 |
| 6.1.2..... | Les |
| <u>divergences</u> | 35 |
| 6.2. LES ENSEIGNEMENTS A PARTIR DES POINTS FORTS ET FAIBLES..... | 35 |
| 7. LE SUIVI DE LA COHERENCE DES DONNEES | 36 |
| 7.1. LES OUTILS D'EVALUATION A METTRE EN PLACE..... | 36 |
| 7.2. L'OBJET DE CES BILANS ET LEURS DESTINATAIRES..... | 36 |
| CONCLUSION | 37 |
| ANNEXES 38 | |
| BIBLIOGRAPHIE | 50 |

INTRODUCTION

La Bibliothèque nationale de France dispose depuis octobre 1999 d'un catalogue unique des livres, périodiques, documents sonores, audiovisuels, informatiques et numérisés représentant huit millions de notices bibliographiques. A terme, il intégrera les documents des collections spécialisées conservées à Richelieu. Il sera la pièce maîtresse du Catalogue Collectif de France (CCFr) qui recensera les fonds les plus importants des bibliothèques françaises.

La Bibliothèque nationale de France a le devoir de produire des notices de référence pour les autres bibliothèques. Du sérieux et de l'objectivité scientifique de sa notice dépend en effet l'accessibilité de l'œuvre, puisque bien décrite et bien classée dans les catalogues, elle sera aisément identifiée, cherchée et trouvée par des générations de lecteurs. Cette règle qui prévaut à la rédaction des notices bibliographiques produites en ligne à partir du dépôt légal peut s'appliquer à l'ensemble des notices du catalogue issues de la rétroconversion ou d'autres sources.

Aussi, l'enjeu du contrôle de la qualité du catalogue est d'améliorer l'accès aux collections et donc de rendre un meilleur service au public en solutionnant les problèmes d'identification et de communication des documents aux lecteurs. Assurer la cohérence dans le catalogue c'est faire en sorte que les différentes œuvres d'une même personne ou sur une même personne puissent être répertoriées sous une forme unique de son nom et correspondent aux unités de conservation dans les magasins. Constituer un outil de référence c'est fournir aux professionnels des bibliothèques et des centres de documentation des données conformes aux normes nationales et aux recommandations internationales, assurant la cohérence de leur catalogue et favorisant les échanges.

Le thème de ce projet concerne donc le contrôle de la qualité et de la cohérence de la base bibliographique BN-Opale Plus, le nouveau catalogue de la Bibliothèque nationale de France.

Comment maintenir la cohérence d'une base alimentée de façon hétérogène ? Comment identifier les erreurs qui peuvent y subsister malgré les contrôles déjà existants et surtout, comment les corriger ? Telles sont les questions que suscite le sujet.

Cela implique d'expliquer les raisons du contrôle et d'énumérer les données envisagées. Cela fait, il faudra déterminer les modes de contrôle et enfin les procédures de corrections pour les erreurs qui auront été détectées.

La BnF disposant d'un seul système informatique pour l'ensemble de ses salles de lecture, le public concerné est le grand public¹, notamment depuis que le catalogue est consultable via Internet² et donc accessible depuis mai 1999 par tout type de public et qu'il sera consultable via le protocole Z 039-50 par le Catalogue Collectif de France.

¹ Cf. PION, Romuald. « Lecteurs, qui êtes-vous ? »

² Via l'adresse suivante : <URL : <http://catalogue.bnf.fr>>

1. CADRE FONCTIONNEL DU PROJET

Afin de mieux situer ce projet, il est indispensable de présenter l'établissement concerné, la Bibliothèque nationale de France, et l'environnement de sa mise en œuvre, le Service de Coordination Bibliographique et en son sein le Bureau Qualité du Catalogue.

1.1. SITUATIONS HIERARCHIQUES AU SEIN DE LA BNF

Etablissement public à caractère administratif créé par le décret n° 94-3 du 3 janvier 1994³, la Bibliothèque nationale de France (BnF) résulte de la fusion de la Bibliothèque nationale et de l'Etablissement public de la Bibliothèque de France (EPBdF) chargé de la construction du nouveau bâtiment sur le site de Tolbiac. Placée sous la tutelle du Ministre de la Culture et de la Communication⁴, ses missions, assignées par le décret du 3 janvier 1994 sont :

« de collecter, cataloguer, conserver et enrichir, dans tous les champs de la connaissance, le patrimoine national dont elle a la charge, en particulier le patrimoine de langue française ou relatif à la civilisation française ». A ce titre, elle exerce les missions du dépôt légal et constitue et diffuse la bibliographie française. Elle rassemble et catalogue des collections françaises et étrangères de tous types de documents et participe à l'activité scientifique nationale et internationale.

Assurer l'accès du plus grand nombre aux collections. A ce titre, elle conduit des programmes de recherche, en particulier sur la bibliothéconomie. Elle coopère avec d'autres établissements, notamment dans le cadre des réseaux documentaires. Elle participe à la mise en commun des ressources documentaires des bibliothèques françaises. Elle permet la consultation à distance en utilisant les technologies les plus modernes de la transmission des données. Elle mène toutes actions pour mettre en valeur ses collections.

L'année 1997 a été marquée par la refonte complète de l'organigramme qui a abouti le 1^{er} octobre 1998 à un regroupement des tâches en 3 grandes directions⁵ :

Direction des Services et des Réseaux
Direction des Collections
Direction des Services Informatiques
Et des Délégations (à la stratégie, aux relations internationales, à la communication).

Etant affecté à la Direction des Services et des Réseaux, la description se limitera à cette direction, et en son sein, uniquement au Service de Coordination Bibliographique.

• *La Direction des Services et des Réseaux (DSR)*

La DSR (ex-Direction du Développement Scientifique et des Réseaux) est divisée en 5 départements et 2 missions rattachées à son directeur. La DSR « est une direction scientifique » qui « contribue à la stratégie générale de l'établissement et notamment aux grands objectifs suivants :

³ Décret n° 94-3 du 3 janvier 1994, Journal officiel de la République française

⁴ Décret n° 83-226 du 22 mars 1983

élargir l'accès aux publics en assurant un service de qualité, fournir des services à distance au plan national et international, développer les relations avec les autres institutions en France et dans le monde»⁶.

La direction a en charge une mission transversale de mise en cohérence et d'harmonisation de la politique scientifique de l'établissement. Elle gère le dépôt légal, les produits bibliographiques, la mise en place et la coordination des réseaux, la politique de la recherche et les échanges internationaux.

L'établissement lui enjoint par là de remplir certaines missions propres à atteindre ces objectifs et pour ce qui concerne le SCB (et ce sujet de PPP), notamment une «amélioration de la qualité et de la disponibilité des produits et services du département du Dépôt légal et de l'Agence Bibliographique Nationale » et une « amélioration de la qualité de service des systèmes d'information »⁷.

- *L'Agence Bibliographique Nationale (ABN)*⁸

« Est chargée de la coordination générale du Dépôt légal. [...] Elle est chargée de l'ensemble des responsabilités scientifiques dans les domaines bibliographiques, catalographiques et de la normalisation documentaire. Elle est notamment responsable de la maintenance des catalogues et de la qualité des produits bibliographiques ⁹ de la BnF »¹⁰. En effet, en tant qu'Agence Nationale elle se doit d'appliquer les deux grands programmes internationaux : le Contrôle Bibliographique Universel (CBU) et l'Accès Universel aux Publications (UAP) pour rendre les documents accessibles.

- *Le Service de la Coordination Bibliographique (SCB, ex-CCBT¹¹)*

Le SCB a été créé en 1985 lors de la mise en place de la base bibliographique BN-Opale (catalogage des documents imprimés). C'est un service transversal qui a pour vocation de coordonner toutes les équipes de catalogage, qu'elles appartiennent à la Direction des Collections (DCO), à l'Agence Bibliographique Nationale ou même à des bibliothèques universitaires travaillant en réseau dans la base BN-Opale. Le SCB¹² est responsable de la cohérence du catalogue et en assure l'intégrité logique ; sa mission principale est de garantir la qualité des données et de l'information bibliographique pour un meilleur service aux lecteurs de la BnF et au réseau des bibliothèques françaises et étrangères.

Il a vocation à étendre ses compétences à toutes les bases bibliographiques en ligne qui alimenteront le catalogage sur le Système

⁵ Cf. « Visage de la BnF [: organigramme] »

⁶ Réorganisation générale des services, 1^{er} octobre 1998 : [document interne]

⁷ Cf. note 6, supra.

⁸ BOUDET, Isabelle. « Agence bibliographique nationale française : tradition et innovation »

⁹ C'est moi qui souligne.

¹⁰ Cf. note 6, supra.

¹¹ Cf. FINELLI-LEMELLE, Françoise. « Le Centre de coordination bibliographique et technique »

¹² Cf. organigramme du SCB : annexe 1

d'Information (SI¹³) : BN-Opaline (documents spécialisés) et Ever Loris (acquisition et catalogage des documents audiovisuels en libre-accès).

- *Le Bureau Qualité du Catalogue (BQC)*

Le 1^{er} juillet 1999 le Bureau de la Cohérence de la Base du Service de Coordination Bibliographique, qui assurait la maintenance de la base BN-Opale a fusionné¹⁴ avec le Service des Chantiers de Conversion Rétrospective, pour devenir une des composantes du SCB, le Bureau Qualité du Catalogue, en intégrant les fonctions et agents de l'ancienne cellule. Les missions de la cellule ont été étendues à toutes les données figurant dans le catalogue et antérieures à la date de démarrage du catalogue en ligne (1980). Il s'agit donc du service de maintenance intellectuelle de la base bibliographique.

Sa mission est d'améliorer globalement et méthodiquement la qualité des données tout en répondant très rapidement aux demandes d'amélioration ponctuelles.

Fondé en novembre 1988, sous le nom de Cèdre (Cellule De RETroconversion) et rattaché au Centre de Coordination Bibliographique et Technique, le Service des Chantiers de la Conversion Rétrospective¹⁵ avait pour mission de convertir les anciens catalogues des anciens départements des imprimés et des périodiques.

Cette phase de travaux s'est achevée en mai 1999 pour l'essentiel (suivi de la saisie, traitement des anomalies, travaux de dédoublonnage internes au marché etc.) et totalement à l'été 1999 (recettes terminées).

1.2. LE PROJET DANS LES OBJECTIFS GENERAUX DE LA BNF

D'après le document d'orientation présenté au conseil d'administration du 24 juin 1998¹⁶, la Bibliothèque doit assurer « un catalogage rapide et efficient de [ses] collections, pour en faciliter l'accès sur place et à distance » et assurer « une politique dynamique de qualité de service [...] ». C'est dans ce cadre que doit s'inscrire l'action du Bureau Qualité du Catalogue : améliorer le catalogage effectué, assurer la qualité des données afin de faciliter l'accès aux documents.

Les travaux sur la base bibliographique sont donc indispensables à deux missions essentielles de la BnF :

- sa mission bibliographique,
- sa mission de communication du patrimoine.

2. MISE EN PLACE D'UNE POLITIQUE DE COHERENCE DES DONNEES RETROCONVERTIES

2.1. LES RAISONS DU CONTROLE QUALITE DES NOTICES

Toute base de données induit l'existence de procédures de mise à jour et de mise à niveau. Le but des travaux à mener est double :

¹³ Il s'agit du système intégré de gestion de la bibliothèque (ou Delisle) dont la partie catalogue se nomme BN-Opale PLUS. Pour plus de renseignements sur le SI, on consultera « Le système d'information de la Bibliothèque nationale de France », *Trajectoire*, juin 1998, n° spécial

¹⁴ fusion effective au 1^{er} juillet 1999 mais qui sera officialisée lors du Comité Technique Paritaire d'octobre.

¹⁵ Devenu un service autonome de la DDSR en 1994.

¹⁶ « Bibliothèque nationale de France : stratégie pour un nouveau millénaire » in *Réorganisation générale des services, 1^{er} octobre 1998 : [document interne]*

- rendre accessible la description des ouvrages, faire connaître leur existence,
- rendre accessibles automatiquement les ouvrages eux-mêmes.

Cette politique de cohérence de la base est nécessaire pour plusieurs raisons.

2.1.1. Une exigence de qualité

Selon la norme NF X 50-120 - ISO 8402¹⁷, la qualité est l'« aptitude d'un produit ou d'un service à rendre les besoins exprimés ou implicites des utilisateurs ». Aussi, l'exigence de qualité est liée au service rendu : contrôler et améliorer la qualité du catalogue c'est rendre service aux usagers.

2.1.2. L'amélioration des résultats de la recherche

Il est des données qui existent dans le catalogue mais qui sont inefficaces pour l'accès automatique aux documents à cause d'erreurs notamment de cotes. Les erreurs notamment de graphie et les lacunes au sein des descriptions bibliographiques ont des conséquences dommageables pour la recherche. Aussi, il convient d'améliorer les accès et la localisation, d'intervenir sur les différentes formes des noms d'auteur et des vedettes en général, gérer les fichiers d'autorité, veiller à l'orthographe des zones titres et des accès supplémentaires (formes translittérées, faux titres, titres de formes et uniformes), posséder une description précise permettant l'identification sans équivoque du document (distinction des rééditions...).

Une erreur sur l'auteur est plus dommageable en recherche par mots-clés qu'en recherche par feuilletage de l'index des auteurs. En effet dans la première méthode la notice dans laquelle se trouve l'erreur sera systématiquement écartée des résultats, tandis qu'en recherche par feuilletage d'index le nom de l'auteur est déclassé à cause de l'erreur, le déclassé est plus ou moins grand selon la place de l'erreur dans le nom. Le lecteur a des chances de la repérer en naviguant plus ou moins loin dans l'index.

Une erreur sur le titre a des conséquences diverses. En recherche par feuilletage de l'index titres le handicap est d'autant plus grand que l'erreur se situe plus près du début du titre : le déclassé est d'autant plus grand. Un balayage est nécessaire, mais il peut être impossible quand il y a un trop grand nombre de titres semblables. En recherche par mots clés la gravité des erreurs est liée aux mots sur lesquels elles portent. Les erreurs les plus graves sont celles sur les mots signifiants du titre, ceux qui sont susceptibles d'être choisis pour bâtir l'équation de recherche (noms propres par exemple, mais pas seulement). Ainsi la notice ayant pour titre « Sur l'empoisonnement des lacs de barrage » est introuvable par mot clé alors que le mot clef est saisi « empoisonnement »¹⁸. Cela est d'autant plus dommageable que ce type de recherche est le plus usité pour les fonds anciens de la BnF puisque ceux-ci ne disposent pas d'indexation matière¹⁹.

¹⁷ *Management de la qualité et assurance de la qualité - Vocabulaire*

¹⁸ Cas réel survenu pour le SU 0031 0482 00 01

¹⁹ L'indexation matière n'est apparue qu'en 1882. Cependant pour relativiser ces propos, il faut savoir que 499000 notices issues de la conversion rétrospective

Notons que les erreurs sur les autres zones de la description bibliographique (qu'auteurs et titres) sont moins dommageables, mais non sans conséquences. Une erreur sur l'adresse (nom de lieu erroné, date de publication fausse), introduit un doute si une édition précise est recherchée. De même, à un moindre degré, une erreur sur la description matérielle de l'ouvrage (format, nombre de pages...). Ces types de renseignements sont toutefois précieux pour tous les bibliophiles et les bibliologues, les spécialistes de l'édition. Les recherches de type historique ou culturel mettant en jeu les lieux de publication (zone adresse, zone fixe positions code pays), les dates de publications (zone fixe positions dates de publication), la langue de l'ouvrage (zone fixe positions code langue) sont incomplètes à la mesure des erreurs commises sur ces éléments ou du défaut de renseignement de ces éléments.

2.1.3. L'uniformisation des données qui servent de référence aux bibliothèques

Pour ce faire il convient de repérer les doublons pour faciliter la consultation et une meilleure gestion des exemplaires et veiller à la bonne gestion des liens entre notices (bibliographique, exemplaire, autorité).

Les doublons sont à éviter et à supprimer car ils encombrant la base en terme de capacité de stockage et de temps de traitement. Les index s'en trouvent surchargés et par là les temps de traitement sont rallongés lors de la consultation de la base, des opérations de maintenance, de nettoyage, etc. Le dédoublonnage ne peut que faciliter les modes de consultation mais aussi de production de la base rendant le choix des catalogueurs moins problématiques notamment pour les formes des autorités.

2.1.4. L'obtention de produits d'édition corrects

Les informations du catalogue sont visibles dans l'OPAC²⁰, la partie catalogue du SI et ce d'autant plus qu'il est consultable via Internet et bientôt par le protocole Z 039 50. Mais elles génèrent en outre des produits puisqu'elles sont éditées sur divers supports (papier, cédérom, disquettes ; cf. le dépliant intitulé *Produits bibliographiques* disponible à l'ABN).

Aussi, s'il est des erreurs de formatage (données saisies dans une zone qui ne leur est pas destinée), les données ne se trouvent ni dans les bons index, ni dans les bons "dictionnaires" de mots clés où on s'attend à les trouver et les notices affectées échappent à la recherche. Ainsi un début de titre saisi en zone auteur rend impossible la recherche de ce titre en mode feuilletage, et souvent en mode mot clés. De même, des erreurs ou l'absence de renseignements dans les zones codées interdisent les tris et nuisent à la recherche. Ces erreurs altèrent la qualité d'éditions thématiques si des tris sont effectués selon des profils car ils sont gênés et incomplets du fait de ces erreurs (ex. virtuel d'un cédérom des éditions parisiennes du XVIIe s.). De même il est des données qui manquent totalement, des notices à convertir ou cataloguer afin de tendre encore plus à l'exhaustivité ou du moins à la complétude du catalogue.

possèdent une telle indexation ce qui est un enrichissement dû à la conversion rétrospective.

²⁰ On line public access catalog

2.1.5. La préservation de l'avenir

Le catalogue ne va pas cesser de croître par le catalogage en ligne des ouvrages reçus au titre du dépôt légal. Le catalogage s'effectuant dans un grand nombre de services répartis entre plusieurs directions les divergences et erreurs, absolument inévitables, doivent être pistées en permanence si l'on veut maintenir la cohérence indispensable pour une bonne qualité de service.

De plus, afin de se prévenir contre le vieillissement des matériels et l'évolution des logiciels, il convient de préserver l'avenir en veillant au format des données qui doit permettre une récupération facile (découpage des informations), au codage des caractères et au langage informatique d'écriture. L'emploi du format MARC utilisant les normes ISO sur les codages de caractères (ISO 646, 5426), sur la structure des données et leur présentation pour les échanges (ISO 2709) permet de résoudre ces exigences. De même la récupération des fichiers nécessite le respect du format, de ne pas créer des zones non prévues et un contrôle de la cohérence du format (champ ou sous-champ répétable ou non, contrôle inter-zones...).

Toute erreur sur les données ou le formatage, tout manque de données ou de liens entre données, est une source potentielle de nouvelles erreurs, pour le catalogage en ligne. Le catalogue est en effet un outil de production de données bibliographiques. Pour assurer la bonne qualité des notices à venir, il convient d'assurer celle des notices existantes.

Mais si les procédures de contrôle de BN-Opale Plus sont à mettre en place, les données qui y figurent ont déjà fait l'objet de vérifications.

2.2. LES CONTROLES DE LA QUALITE DEJA EFFECTUEES SUR LES DONNEES DU CATALOGUE

2.2.1. Lors du catalogage en ligne

La base BN-Opale, production d'une agence bibliographique nationale est de bonne qualité par l'application stricte des normes, du choix des accès et la bonne gestion des autorités. Elle fait l'objet d'une maintenance intellectuelle assurée par la section "Cohérence de la Base" du Service de Coordination Bibliographique.

La taille de la base BN-Opale et l'importance croissante de la production éditoriale interdisent un contrôle systématique livre en main du travail des catalogueurs. Cependant, des contrôles de la base sont menés : d'une part des contrôles en ligne lors de la saisie des notices qui font apparaître des messages d'erreurs au catalogueur et d'autre part des vérifications a posteriori de deux types : d'une part en ligne²¹ ; d'autre part, tous les quinze jours une relecture complète d'un listing des notices en format InterMarc produites pour fournir le bon à tirer de l'édition papier de la Bibliographie Nationale Française.

De même, les données qui figurent dans BN-Opale Plus issues de la rétroconversion ont subis des contrôles.

²¹ Les contrôles en ligne existants ou qui seront mis en place dans le SI ne sont pas davantage explicités n'étant pas l'objet de ce projet.

2.2.2. Lors de la conversion rétrospective

Une conversion rétrospective de masse²² (notamment pour des volumes tels ceux de la BnF) interdit une relecture totale de chaque notice saisie. Cependant des travaux ont été menés avant la saisie en masse pour minimiser les risques d'erreurs : préparation de catalogues et fichiers (parfois intégralement reformatés en InterMarc), travaux préparatoires et formation des prestataires, rédaction de référentiels pour faciliter la saisie (règles de découpage des zones et sous-zones, règles d'uniformisation des cotes, des codes pays, constitution d'échantillon test de saisie).

Le travail ayant été sous-traité auprès de prestataires de service extérieurs à la BnF, il se devait d'être d'autant plus contrôlé. Le but de la conversion était d'uniformiser au maximum les données converties afin d'éviter au maximum les erreurs impliquant une reprise des notices, une correction a posteriori. Aussi, le traitement de chaque lot de saisie de conversion rétrospective a compris des phases de contrôles. D'une part au cours de la saisie le prestataire a le droit de réclamer sous forme d'anomalie des précisions sur les données afin de les saisir au mieux. Ainsi au cours de l'opération sur une saisie de 4 429 089 notices, 161 000 ont fait l'objet d'anomalies. Elles ont donc été entièrement relues par des bibliothécaires et peuvent donc être considérées comme sans erreur grave et conformes, non par rapport aux normes les plus récentes, mais par rapport au fichier d'origine. De plus, grâce au lot 6 (dédoublonnages) : 76 411 notices ont été contrôlées par la BnF. Il y a donc eu : 187 134 notices passées au contrôle BnF, soit un pourcentage de : 4,85 sur l'ensemble de la phase 1. D'autre part, les contrôles qualité mensuels ont permis la vérification, par un prestataire, de : 127 500 notices, pour lesquelles les fautes les plus graves ont pu être corrigées dans le cadre du lot 6. Pour la phase 2, sur 326 945 notices saisies, 1751 ont été supprimées et sur 67 707 reçues pour contrôle 24 901 ont été relues. Il reste 42 806 notices à relire. D'autre part des contrôles de trois types sont menés en cours de saisie :

un contrôle du format de la bande et de la séquence des identifiants : la bande informatique qui stocke les notices a été contrôlée dans sa structure pour ne pas rendre les données inexploitable d'un point de vue informatique et les identifiants ont été contrôlés dans leur séquence.

un contrôle d'exhaustivité : le contrôleur, qui est un prestataire extérieur différent du prestataire de saisie, procède à deux opérations :

saisie des seuls identifiants de l'échantillon de contrôle
comparaison au fichier des identifiants saisis par le prestataire de saisie pour en relever les différences : identifiants en trop, différents ou manquants et contrôle de la structure des identifiants,

²² Sur la conversion rétrospective à la BnF, on consultera pour en connaître les principes et l'historique : BEAUDIQUEZ, Marcelle. « La conversion rétrospective en chiffres » ; BEAUDIQUEZ, Marcelle. « La conversion rétrospective : geste historique au service de la mémoire et mission de service public, la conversion rétrospective est aussi l'œuvre d'une équipe » ; BEAUDIQUEZ, Marcelle. « Le chantier conversion rétrospective de la Bibliothèque nationale » ; BEAUGENDRE, Anne-Marie, EHLING, Thérèse. « La conversion rétrospective à la portée de tous » ; BERNARD, Annick. « Conversion rétrospective - Conversion « autonome » : d'après l'expérience de la Bibliothèque nationale » ; SOLOMIAC, Jacqueline. « La conversion rétrospective à la Bibliothèque nationale ».

repérage des identifiants erronés, des identifiants en anomalies, rupture de séquence, comptage, statistiques par types de notices.

un contrôle qualité : un échantillon témoin de comparaison est élaboré d'après la norme NF X 06-022 qui définit la taille de l'échantillon à contrôler variable selon le nombre total de notices du lot, la répartition par type de notice et le niveau de qualité acceptable (NQA) qui précise le taux de non conformité en nombre d'individus et en nombre moyen de caractères non conformes par 100 unités pour chaque sous-lot. Des grilles de contrôles présentant des niveaux de qualité modulables en fonction de la qualité initiale de chaque fichier ont été établies pour déterminer les zones à risques et les zones aux erreurs moins «gênantes». Le prestataire extrait les notices à saisir, les balise et les saisit pour constituer un fichier témoin qui est ensuite comparé avec la saisie du prestataire de saisie pour fournir en fin, des statistiques par type d'erreur et un listing confrontant les deux saisies, la sienne et celle du prestataire de saisie pour chaque notice où il a repéré une différence de traitement.

De plus, des contrôles automatisés sont menés lors de la saisie quant au contenu des notices : sur les zones (obligatoire ou interdite), sur la structure (zone numérique, date...) et sur la codification (par rapport à des tables de valeurs).

La «recette» d'un lot, c'est-à-dire la validation par la Bibliothèque nationale de France intervient après analyse de ces contrôles. La facturation d'un lot n'intervient qu'en cas de recette positive car sinon, le prestataire de saisie doit recycler le lot à ses frais, c'est-à-dire le ressaisir entièrement pour recommencer la procédure des contrôles.

Ces contrôles sont clos une fois le fichier validé, c'est-à-dire recetté positivement, accepté par la BnF. La recette d'un lot de saisie de conversion rétrospective n'intervient qu'après contrôle. Le taux d'erreurs maximum possible est connu de façon détaillée, catalogue par catalogue²³, puisqu'il était défini dans chaque cahier des charges et conditionnait la recette positive des sous-lots mensuels de saisie. Ce taux avait été adapté à la difficulté de décryptage de chaque fichier ou catalogue et avait été approuvé lors des comités de pilotage. Le niveau de qualité requise (NQA) auquel il aboutissait avait même été jugé à l'époque très élevé, voir même impossible à atteindre par l'une des sociétés soumissionnaires. Cependant les différents NQA exigés ont été généralement atteints, parfois n'ont été qu'approchés pour des fichiers très difficiles (AN), et au contraire ont parfois été nettement dépassés (CG). Il est donc valable de considérer que le nombre d'erreurs produites par la saisie est approchant du nombre qui avait été jugé comme un maximum à tolérer.

Enfin, après la conversion des travaux ont été menés afin d'améliorer la qualité des données : dédoublonnage de fichiers, uniformisation des accès, autorités créées automatiquement.

2.3. LE CONTROLE DE LA QUALITE A MENER SUR LE CATALOGUE BN-OPALE PLUS

Le contrôle qualité ne doit pas cesser sous prétexte que des contrôles ont été menés lors de la saisie de la conversion rétrospective ou que des contrôles sont effectués par le Dépôt légal. L'élargissement du public

²³ Cf. annexe 2.

concerné par le catalogue informatisé ainsi que les volumes importants de son contenu induiront une demande accrue de sa maintenance. Aussi, comme pour toute base de données importante, des contrôles et des corrections s'imposent pour en assurer la cohérence. La démarche qualité doit aussi bien s'appliquer en amont²⁴, à la production des notices qu'en aval, au produit fini.

Car en effet, la maintenance des données proprement bibliographiques (c'est-à-dire hors fichiers d'autorité) va faire un saut quantitatif considérable avec le déploiement du SI : on va passer d'un coup de moins de 2 M de notices à plus de 6,5 M en version 1 (V1) du Système Informatique (SI), 7 M en version 2 (V2), 7,2 M après les derniers chargements de la phase 2 de la Conversion rétrospective des imprimés, 8,2 M quand la conversion des départements spécialisés sera achevée (pour les imprimés seulement).

Un service de maintenance doit être d'autant plus permanent qu'aux erreurs présentes à la constitution de la base viendront s'ajouter celles faites au cours du travail courant en ligne par les catalogueurs²⁵. Aussi, BN-Opale Plus nécessite deux grands types de corrections :

des améliorations urgentes à mener en priorité notamment sur les données rétroconverties,

et une maintenance intellectuelle au long cours²⁶.

Et ce d'autant plus que la publicité des erreurs va s'accroître du fait de la mise à disposition pour la première fois de l'ensemble des catalogues de la BnF en un seul catalogue. Cela va mettre en exergue les incohérences des anciens fichiers, qui étaient moins frappantes auparavant. Il en va de même par la mise en réseau, le déploiement général de ce catalogue sur l'ensemble des sites de la BnF et à l'extérieur qui va entraîner un accroissement des remarques critiques et des demandes de corrections, de la part des professionnels, des lecteurs et du public non connaisseur de l'histoire de la BnF et de la constitution du catalogue qui le consulte de l'extérieur de la BnF.

Pour toutes ces raisons, différentes opérations sont à mener sur les données pour offrir un catalogue convenable : d'une part identifier les problèmes et d'autre part, les régler. Cela suppose une première phase d'identification et de recensement des problèmes puis une deuxième phase de correction.

3. L'IDENTIFICATION DES PROBLEMES

Afin de mener un programme de corrections méthodiques, il convient de réaliser un état des lieux aussi exhaustif que possible des erreurs avant d'établir des priorités de traitement dans les procédures d'amélioration.

²⁴ La qualité du catalogue concerne l'ensemble des notices, celles qui sont versées comme celles qui sont créées quotidiennement. L'objet de ce PPP concerne celles qui sont versées mais bien entendu il convient de continuer la formation des catalogueurs, de contrôler leur travail et notamment de vérifier la précision des données, le respect de la codification, la cohérence du fichier, la présence des formes de renvoi dans les autorités, de résoudre les problèmes d'homonymie, de repérer les doublons et veiller à l'application des règles de catalogage et des normes.

²⁵ La production courante risque de créer des doublons lorsque trop d'autorités élémentaires répondent à leur recherche auteur ou en se liant de façon aléatoire à une entrée d'index puisque les impératifs de production ne vont pas leur permettre des recherches longues dans toutes les entrées.

²⁶ Les contrôles du catalogage courant en ligne à venir ne sont pas l'objet de ce projet. Ils sont prévus mais le marché ayant été résilié, il paraît plus qu'aléatoire de les évoquer.

Avant d'agir sur des notices, de les corriger, il convient de les identifier, de repérer celles qu'il faut traiter. Cela semble être une évidence mais l'identification est ce qui pose le plus de problèmes. Car en effet, comment repère-t-on les notices fautives, comment identifie-t-on dans une base de 8 millions de notices celles qui comportent une erreur ? La littérature professionnelle est à ce propos assez pauvre privilégiant l'aspect dédoublonnage, facile à programmer, au repérage de notices isolées et uniques.

3.1. LES ERREURS A CORRIGER DEJA IDENTIFIEES ET LEUR VOLUME

BN-Opale Plus est une base hétérogène puisqu'elle résulte de la fusion de BN-Opale et de Geac libre-accès, la base qui recense les ouvrages acquis pour le libre-accès de Tolbiac, de chargements des notices rétroconverties, et de chargements des cotes issues de la base Sycomore²⁷. A ce titre la qualité des notices qui la composent est inégale puisque les notices de BN-Opale sont conçues d'après les normes les plus récentes de catalogage, que celles de Geac-Ivry répondent à la norme de catalogage allégé et que celles de la conversion rétrospective sont les plus diverses dans leur présentation puisque certaines ont été créées d'après les « cartes à jouer » commencées sous les injonctions de l'abbé Grégoire.

Les sources d'erreur sont multiples et varient suivant leur origine : la plus grande part provient de la rétroconversion mais il en est issues de BN-Opale ou de Sycomore.

3.1.1. Les erreurs de la rétroconversion

Les contrôles devront porter sur l'ensemble de la base mais en priorité sur les notices en provenance de la conversion rétrospective des imprimés. Car la part des notices issues de ce chantier est la plus importante dans le catalogue et la plus sujette à erreurs.

Les erreurs dans les notices rétroconverties de BN-Opale Plus proviennent soit des catalogues d'origine, soit sont dues à la saisie.

- *Des erreurs originales*

Les erreurs sont nombreuses dans les catalogues d'origine, y compris les catalogues imprimés de la BN, et elles ont été transcrites telles quelles lors de la saisie. Les sociétés prestataires n'étaient pas tenues de porter un regard critique sur les données à saisir, ce qui d'ailleurs leur aurait été dans la majorité des cas impossible. D'ailleurs bon nombre de ces erreurs n'apparaissent absolument pas à la seule vue du catalogue (cote fausse mais vraisemblable, nom propre mal orthographié...). En effet, les catalogues et fichiers convertis n'étaient pas homogènes. Rares furent les catalogues et fichiers comportant des notices bibliographiques conformes aux normes de catalogage actuelles quand elles ne se bornaient pas à être de simples fiches d'inventaire ne comportant quasiment aucun élément du pavé ISBD (ex. fichier

²⁷ SYCOMORE (système de communication des ouvrages récolés) est la base où sont enregistrées les cotes des documents qui ont été récolés et conservés dans les magasins et salles de travail de Richelieu. Elle gérait la communication aux lecteurs des livres et périodiques. Le signalement des exemplaires recensés a été rapproché des notices bibliographiques pour constituer le catalogue informatisé unique.

des ventes d'art). Les catalogues avaient des présentations d'une qualité et d'une cohérence inégales dans leur logique de classement (auteur, matière, cote), leur présentation intellectuelle (rééditions présentées de façon abrégée sous une édition originale, collections avec ou sans dépouillement), ou physique (fichiers manuscrits, imprimés ou dactylographiés ; catalogues imprimés ou autographiés). Toutes ces variétés de présentation peuvent se trouver mélangées dans un seul catalogue. C'est pourquoi une conversion implique la mise en place de règles de saisie pour aboutir à un format logique. Pour aboutir à un catalogue informatisé unique respectant un format cohérent à partir de ces sources peu normalisées, des règles d'uniformisation strictes ont été mises en place pour chaque sous-lot en fonction de sa spécificité propre. Aussi les cahiers des charges fournis aux prestataires étaient-ils denses et volumineux et le travail fourni aux prestataires extérieurs a été contrôlé pour garantir la qualité des données obtenues.

Mais ce phénomène d'hétérogénéité, général dans les catalogues anciens rédigés avant l'apparition et l'application des normes nationales (apparues peu avant 1970, rappelons-le alors que les catalogues convertis sont tous antérieurs à cette date) a provoqué des erreurs de saisie.

- *Des erreurs de saisie ou de code*

Malgré une saisie en double frappe demandée aux prestataires, il existe des erreurs dues à la saisie proprement dite. Le temps ayant manqué pour préparer tous les fichiers à convertir, les travaux de tri préalables, de barrage des informations inutiles, de regroupement d'exemplaires catalogués sur des fiches différentes, d'uniformisation des points d'accès n'ont pu être menés à terme. Cela a pu entraîner des erreurs de déchiffrement et de transcription lors de la saisie, de la frappe des données, ou des erreurs de formatage : les données n'ont pas été saisies dans la zone adéquate du format informatique et ne peuvent donc pas être utilisées pour la recherche. Ces erreurs peuvent persister dans la base car lors des contrôles qualité des sous-lots de saisie (cf. point 2.2.2.), des tableaux d'erreurs inacceptables entraînant le rejet du lot (son rejet en cas de recette négative ou son acceptation) sont renseignés mais aucune correction n'est apportée aux notices. Il faut bien comprendre que des erreurs sont repérées mais pas corrigées. Seul un recyclage peut laisser espérer qu'une correction interviendra, c'est-à-dire que le prestataire de saisie la repérera pour la corriger lors de la 2^e saisie complète du lot. Sinon, pour les lots acceptés, l'erreur persiste dans la bande livrée et donc chargée dans le SI malgré les contrôles menés.

Ainsi, les notices comportant des caractères non latins ne pouvant être convertis avec les jeux de caractère (de la norme ISO 5426-1983) possèdent à la place des dits caractères le signe « @ ». Même s'il ne s'agit pas à proprement parler d'une erreur, cela constitue autant de notices à reprendre surtout quand ce symbole pollue les index, quand il concerne un champ d'accès (titre, auteur...). Environ 90 000 notices sont en caractères non latins mais accessibles par les auteurs en latin et les cotes.

Les zones codées et notamment la 008 comportent des erreurs :

- les notices ne possédant pas de date en zone adresse (260 \$ds.d.) se retrouvent avec une position date en 008 à blanc,
- des codes ont changé (évolution de la liste internationale ISO des pays),
- les notices comportant le code pays XX, le code langue UND ou la position date non renseignée peuvent faire l'objet d'une amélioration,
- des erreurs de codes ont été saisies par les prestataires (ex. : ve Venezuela pour vn Vietnam ; confusion entre Guinée-Bissau et Guinée-Equatoriale, entre Congo belge rw, Congo français ga et Congo cg),
- des zones adresses et collation n'ont pas été différenciées à la saisie dans certains catalogues (absence de zone 280 dans les notices du SU),
- des notices (TD/TL) possèdent la ville de soutenance en 710 ce qui est à enlever.

- *Des doublons*

Il y a doublon dans un catalogue informatisé quand il existe plusieurs notices pour un même ouvrage, plusieurs descriptions pour une seule entité physique. Les doublons du catalogue proviennent soit de notices secondaires qui n'ont pas été fusionnées avec les notices principales, soit différents exemplaires possédés par la bibliothèque ont été catalogués chacun séparément. Ils "auraient" du disparaître lors des différents programmes de travaux automatiques de dédoublonnage²⁸ déjà effectués soit dans certains marchés de conversion, soit avant la migration dans la V1, soit à la migration en V1. Si des doublons persistent cela n'est pas du à des erreurs dans la conversion ni dans les programmes de dédoublonnage, mais à l'incohérence du catalogage, à son hétérogénéité qui entraîne des doublons de notices : le catalogue compte en effet des documents de types différents, au format plus ou moins complet, et selon des règles de catalogage d'inégale cohérence. L'absence de norme, et notamment pour les vedettes, entraîne des redondances intellectuelles de zones dans les notices résultantes et donc des doublons qui n'ont pas été dédoublonnés lors des divers traitements subis par les notices car non strictement identiques. Ces doublons sont de deux types :

- des notices plus ou moins identiques recouvrant la même UC ("unités de conservations" : volumes, ou parties de volume constituant une unité bibliographique) : il s'agit de 2 notices appelées à n'en devenir qu'une avec possibilité d'enrichissement de l'une par l'autre. La plupart de ces notices, mais pas la totalité, sont des notices doublons,
- des doublons des notices d'autorité. C'est le point d'accès autorité personne physique qu'il faut uniformiser d'autant que les fichiers d'autorité de la BnF serviront de base au référentiel d'autorités du SU, soit environ 550 000 notices d'autorité. Il faut rappeler que lors de la constitution du SI, le nombre des autorités a interdit une uniformisation manuelle de chacune. Aussi, il a été procédé à des regroupements de notices bibliographiques possédant des autorités semblables mais pas forcément identiques sous une autorité élémentaire. On obtient in fine une grappe de notices sous une autorité²⁹. Il convient donc de distinguer les homonymes, par

²⁸ Sur la nécessité de dédoubler on consultera : DESRICHARD, Yves. « Le dédoublonnage des banques de données bibliographiques : un état de l'art » et les articles anglo-saxons d'identification des doublons.

²⁹ Cf. sortie imprimée du SI de la grappe d'autorité « Balzac » en annexe 6.

leurs dates biographiques et/ou leur profession ; attribuer à chaque auteur les œuvres auxquelles il a participé et uniquement celles-ci.

- enfin, il y a 22 000 notices avec une zone 983 qui indiquent qu'elles sont le doublon d'une ou plusieurs autres notices dont on a noté l'identifiant dans cette zone,

Cela constitue autant de doublons repérés mais qu'il faudra traiter à la main. Pour repérer d'autres doublons, il faudra commencer par feuilleter les index titres. L'uniformisation des autorités personnes physiques devrait permettre de repérer nombre de doublons subsistants. Enfin, il faudra recourir aux clefs de recherche citées abondamment par la littérature anglosaxonne (cf. Bibliographie, 3^e partie).

- *Des notices à erreurs multiples*

Le BQC dispose déjà de notices à traiter et pour lesquelles il faut faire des recherches (BN-Opale Plus, BN-Opale, fichiers et catalogues papier, base Sycomore et magasins) afin de les corriger et améliorer. Elles peuvent receler des erreurs de cotes, des doublons et autres. Il s'agit :

- des lots dits de « retards » composés des anomalies qui n'ont pu être corrigées, faute de temps, par la BnF et qui peuvent comprendre tous les types d'erreurs et nécessitent par là une relecture complète (AN = Anonymes avant 1960, Factums (FH, FA), thèses de médecine (TM), liturgie (LI)). Pour les Anonymes il s'agit souvent de vérifier que tous les volumes d'une collection éditoriale ont bien une notice de monographie avant de rectifier la notice de collection. C'est-à-dire qu'il faut éclater les dépouillements en autant de notices principales et donc d'améliorer la notice principale quand elle existe ou la créer. Les liens entre notice de monographie et notice de collection seront faits dans BN-Opale Plus. Un dixième de ces volumes a échappé à tout catalogage particulier mais dispose d'une UC (cote de Sycomore). Pour les thèses de médecine il est bien connu que la quasi-totalité des cotes des thèses soutenues avant 1935 sont à reprendre ou à reconstituer, à l'aide des "Carnets" (registres anciens d'inventaire par cotes) et de SYCOMORE notamment.

- des listings des divers traitements de rapprochements pour dédoublonnages chez les prestataires de saisie lors de la phase 2. Ces listings sont très intéressants à exploiter car ils mettent en évidence la totalité des erreurs sur les cotes, y compris celles des fichiers et catalogues d'origine. Malheureusement seuls certains fichiers sont concernés (à titre d'exemple : Histoire de France 1895-1950 sur fiches, factums, Smith-Lesouëf, catalogues de ventes publiques). Des vérifications en magasin s'imposent avant de pouvoir restituer et corriger les épines dorsales.

3.1.2. Les erreurs de BN-Opale, Sycomore ou dues à la migration

- *Dans les données provenant de BN-Opale*

Même si elle a été constituée principalement en ligne, la base BN-Opale comporte des erreurs qui n'ont pas pu être toutes corrigées avant leur migration dans le SI. D'autres erreurs se sont produites pendant la migration.

Les erreurs repérées portent essentiellement sur le formatage, les liens et les doublonnages. Elles sont répertoriées dans un document du SCB intitulé "Recensement des corrections à effectuer sur le catalogue en version 2 du S.I. (P.C.A.)" d'avril 1998, qui liste déjà plus de 80 types de corrections.

- *Du récolement*

Il existe aussi des erreurs provenant du récolement dues aux erreurs originelles des catalogues papier mais essentiellement à des pratiques différentes de saisie dans Sycomore par rapport à la présentation des cotes des catalogues qui sont des causes de non constitution de l'épine dorsale (attribution de sous-cotes essentiellement). Ainsi parmi les notices qui doublonnent (citées plus haut), celles qui ne sont pas des doublons, et qui ne nécessitent donc aucune modification, résultent du fait que dans certains cas le catalogage a été bien plus fin que le récolement : par exemple, on a catalogué à l'unité les pièces d'un recueil, on a fait le catalogage détaillé de publications en série scientifiques, alors que le récolement n'a pris en compte que des volumes entiers ou des séries entières. Dans ce cas, les multiples notices bibliographiques pointent toutes sur l'une d'entre elles qui est elle-même rattachée à l'unique UC.

Or, dans le SI, on distingue deux parties : le "Haut" constitué des notices catalographiques provenant des fichiers de la CRI (Conversion rétrospective des imprimés) et le "Bas" constitué des données sur l'exemplaire et les unités de conservation qui proviennent de la base de récolement (Sycomore). Pour constituer le SI les notices de la CRI ont été rapprochées des UC pour créer les épines dorsales ce qui parfois n'a pas eu lieu. La non constitution de l'épine dorsale n'est pas toujours due à une mauvaise qualité de la conversion, mais à d'autres causes :

- des cotes fausses : cotes vraisemblables mais n'existant en fait pas,
- des cotes incomplètes (par exemple absence de sous-cote),
- des cotes absentes dans le catalogue (absence totale, ou simple indication de lettrage),
- des ouvrages disparus en magasin et non enregistrés comme tels par le récolement,
- des ouvrages du fonds non encore récolés.

- *De la migration*

Des erreurs sont apparues lors de la migration des données rétroconverties, c'est-à-dire le versement des notices dans le SI effectué après une transcodification du format Interarc en Interarc intégré. En effet, il est par exemple arrivé que des zones existantes dans des fichiers mais n'étant pas prévues dans le format Interarc intégré soient recopiées dans une zone de stockage (142 étiquettes 743 des anonymes mises en 999 à réindexer).

3.2. LES PROBLEMES INCONNUS A IDENTIFIER ET COMMENT

Restent les problèmes non identifiés qu'il faudra bien corriger également. La difficulté réside dans leur repérage. Comme il est impossible d'imaginer

où peut résider l'erreur il est par voie de conséquence impossible de la repérer systématiquement. Le BQC compte sur les remarques des usagers et sur les bordereaux de correction qui lui seront transmis pour tenter de repérer des types d'erreurs systématiques qu'il serait éventuellement possible de pister par un programme informatique, d'éditer sous forme de listing à traiter. L'analyse de ces demandes devra donc être attentive et critique afin de lancer des sondages pour soulever les erreurs récurrentes. De même la participation au service public de l'ensemble des agents du Bureau devrait permettre de repérer des erreurs lors des recherches menées pour les usagers car le meilleur mode de repérage reste la consultation régulière de la base notamment selon des critères imprévus. Le regard extérieur des lecteurs ne pourra qu'être bénéfique car leur usage de la base ne se veut que pratique. C'est bien la recherche que le Bureau souhaite améliorer par le contrôle de la qualité.

Un autre mode de repérage peut consister dans la lecture des index de la base pour toutes les zones faisant l'objet d'un index. Pour celles qui ne sont pas indexées comme les notes où, comme dans toute zone de texte, il est des erreurs en nombre, il pourrait être envisager de demander une édition de ces zones par lots de saisie de conversion. Cela permettra de repérer les erreurs, les champs à contrôler et donc éventuellement d'écrire en langage clair les erreurs à repérer, les donner à un informaticien qui les transformera en langage SQL (Standard Query Language : langage permettant d'interroger des bases de données relationnelles) pour sortir des listings.

3.3. LA VOLUMETRIE DES ERREURS

3.3.1. De la conversion

A partir des NQA donnés dans les différents cahiers des charges on obtient une estimation maximum du nombre d'erreurs qu'on peut trouver dans chaque catalogue (cf. annexe 2). Il est à noter qu'il existe des catalogues où :

- le niveau de qualité est supérieur au NQA : CG, HL
- il y a un peu plus d'erreurs que toléré dans le cahiers des charges : AN.

On peut estimer globalement que ces taux permettent d'obtenir un véritable ordre de grandeur. Ainsi, 517 887 erreurs, toutes espèces confondues au maximum, peuvent exister sur l'ensemble de la CRI (phase 1 et phase 2 complètes). Il est à noter que les erreurs sur les cotes comptabilisées au moment des recettes (dont les premières remontent à la mi-1991 et sont antérieures à l'existence de la base Sycomore) ne sont pas forcément celles qu'on comptabiliserait maintenant ; nous savons depuis que l'on a accès à Sycomore, quelles sont les « vraies » erreurs, celles qui empêchent la constitution de l'épine dorsale. C'est l'une des causes de divergence entre le nombre d'erreurs sur les cotes calculé d'après le NQA (les erreurs par rapport aux catalogues papier sont au maximum de 15 608) et celui calculé d'après les résultats de la migration pour V1 (638 921 qui n'ont pas retrouvé leur équivalent dans Sycomore).

Le volume d'erreurs déduit des NQA ne prend en compte que les erreurs dues à la conversion. Il n'est pas possible de chiffrer le volume des erreurs originelles qui subsistent après la correction d'un certain nombre d'entre

elles lors des opérations de saisie, lors de rapprochements (phase 1 seulement) et lors de l'opération COTRETRO³⁰.

Une UC peut correspondre à des dizaines de notices bibliographiques (par exemple pour un recueil de pièces) et inversement une seule notice bibliographique peut correspondre à des centaines d'UC (périodiques, séries...).

Il est donc, pour toutes ces raisons, très difficile de s'approcher du volume réel des notices ayant besoin d'être retouchées ou détruites : 450 000, 780 000, 1 100 000... même si le plus souvent une seule et même notice cumule plusieurs types d'erreurs, c'est de toutes façons en centaine de milliers qu'il faut évaluer le nombre de notices à retoucher.

Le tableau ci-dessous donne la synthèse du nombre d'erreurs possibles par rapport aux catalogues d'origine pour toute la conversion rétrospective :

| VOLUMETRIE MAXIMUM DES ERREURS DUES A LA CONVERSION | |
|--|---------|
| Cotes | 15 608 |
| Vedettes | 33 978 |
| Autres | 468 301 |
| TOTAL | 517 887 |

- *Les erreurs de cotes*

Le Bureau dispose de 2 moyens pour une approche du nombre de cotes en erreur (cf. annexe 3) :

- les erreurs décelées ayant fait l'objet d'un bordereau de demande de correction : 1726 au 1.7.1998 ; ces erreurs sont le plus souvent des erreurs originelles,
- les résultats de la constitution de l'épine dorsale pour la constitution de la V1 du SI.

Le taux de constitution de cette épine dorsale pour les données issues de la conversion rétrospective (migration pour V1) varie de 94,3 % à 77,8 % (Fonds de Fontainebleau exclu : 26,5 % : les ouvrages encore au château n'ont pas été récolés). A titre de comparaison le taux est de 96 % pour les données issues de BN-OPALE (les cotes des départements spécialisés et les cotes trop aberrantes, non "normalisables" par la migration, ont été exclues pour le calcul de ces taux). Pour ce qui est du "Haut" du SI seul, et pour l'ensemble des catalogues présents en V1 y compris BN-Opale, 638 921 cotes des catalogues n'ont pas trouvé leur équivalent dans le récolement.

Pour ce qui est du "Bas" seul, 1 781 474 UC n'ont pas trouvé de notices. Cela est normal pour les unités de conservation qui correspondent aux 674 000 notices de la conversion rétrospective non disponibles pour le chargement en V1. Mais, à supposer qu'à une notice corresponde une UC (ce qui est inexact dans bien des cas), il resterait encore 1 107 474 UC orphelines de catalogage. Certaines ne se sont pas constituées du fait que :

- soit la notice bibliographique existe mais avec une erreur sur sa cote,
- soit la notice bibliographique n'existe pas : la notice n'est pas dans les catalogues de la conversion ou l'ouvrage n'a jamais été catalogué.

³⁰ Opération menée par le SCCR en 1996-1997 en collaboration avec le service Sycomore, préalablement à la migration dans V1 pour uniformiser le maximum de cotes de la CRI avec la saisie de Sycomore. 50000 cotes ont été corrigées.

Il n'est pas possible de savoir combien des 638 921 "hauts" sont appareillables manuellement, avec ces 1 107 474 "bas", ni combien de notices bibliographiques sont purement et simplement à créer (catalogage). Il n'est pas possible non plus de savoir s'il s'agit de livres anciens (avant 1830) ou d'ouvrages relevant d'un catalogage plus courant.

Une autre approche quantitative consiste à calculer à partir des pourcentages d'épines dorsales constituées et du nombre de cotes présentes dans les catalogues, le nombre de cotes des catalogues qu'il faudra vérifier et tenter d'appareiller aux cotes du récolement (plus précisément aux UC= unités de conservation). Le chiffre provisoire est de plus de 400 000 cotes à vérifier. Ce chiffre montre que l'essentiel des notices n'ayant pas trouvé leur équivalent dans le récolement provient de la CRI et non de BN-Opale, ce qui n'est pas surprenant. Par ailleurs la CRI comprend des notices de fonds qui n'ont pas été récolés (par exemple : Fontainebleau, Smith Lesouëf, Bulletins paroissiaux, etc.).

- *Les doublons*

On peut tenter de les quantifier en connaissant le nombre de notices bibliographiques qui ne sont pas directement liées à une cote issue du récolement mais qui le sont indirectement, par un pointeur vers une autre notice bibliographique portant la même cote. La migration en V1 a recensé 808 675 notices bibliographiques rattachés à leur UC en transitant par une autre notice bibliographique (soit 9,59 % des de V1).

Si ce nombre apparaît comme nuisible à la consultation et est le signe d'une moindre qualité du catalogue il faut savoir que les catalogues papiers étaient bien plus encombrés de doublons du fait de leur structure : lors de la conversion environ 250 000 doublons ont été supprimés.

- *Les autres erreurs*

Outre les doublons et les erreurs de cotes, il est d'autres erreurs comptabilisées :

Les erreurs sur les vedettes sont au maximum de 33 978.

Les erreurs originales ne sont malheureusement pas quantifiables et se découvrent par hasard au fil des années. A titre d'exemple, rien que pour les données de la phase 1 de conversion rétrospective, le SCCR reçoit ou rédige une centaine de demandes de correction chaque mois.

Le Bureau dispose de notices à problème disponibles sur papier et recelant tout type d'erreurs (de cotes, des doublons et autres) au nombre de 100 000 :

- 50 000 notices en anomalies du catalogue AN (Anonymes avant 1960),
- entre 40 et 50 000 notices en anomalie des Factums (FH, FA), thèses de médecine (TM), liturgie (LI)).

- 7 000 corrections sur bordereaux concernant la CRI (conservés au SCCR).

Les autres erreurs se répartissent entre erreurs de frappe dans des zones autres que les points d'accès principaux, mauvais codages de pays ou de langues, mauvais formatages. Elles forment la grosse masse, mais leur impact sur la qualité du catalogue est bien moindre. Il faut noter que les erreurs de formatage ont été bien moins nombreuses, en fait, que les erreurs de graphie ou de codes. Quant aux erreurs dans la zone fixe de codification

(zone 008), elles n'ont pu être décelées par programmes : elles sont vraisemblables selon les tables de références, mais souvent fausses intellectuellement.

3.3.2. Hors conversion

Les erreurs dues à la migration sont peu nombreuses pour la conversion. Par exemple une certaine zone point d'accès aux notices (743 = autre titre de forme), qui n'était pas prévue dans les règles de migration s'est trouvée reversée dans une zone de rejet, qui n'est pas indexée et qui n'est pas visible du public. Une seule notice et quelques dizaines de zones ont été rejetées au test de migration de 1997.

Les erreurs dues au chargement de BN-Opale portent sur quelques dizaines ou centaines de notices, d'autres sur des centaines de mille, mais la plupart ne sont pas quantifiables exactement. Le Bureau dispose d'environ 5 000 bordereaux de correction concernant BN-Opale proprement dit (conservés au SCB).

4. L'AMELIORATION DU CATALOGUE

Le but du service est d'améliorer globalement et méthodiquement la qualité des données et de répondre rapidement aux demandes d'amélioration ponctuelles.

L'intervention se fait a posteriori puisque qu'un recatalogage complet est à exclure faute de temps et de moyens en personnel (pour un recatalogage en interne) ou financier (pour un recatalogage sous-traité).

4.1. LA CORRECTION DES ERREURS : METHODE ET ORGANISATION DU TRAVAIL

Le Bureau doit réfléchir à des interventions manuelles à faire au jour le jour, pour satisfaire la demande des utilisateurs (personnel BnF, public interne et externe) et étudier les possibilités de corrections de masse par programmes informatiques à accomplir dans le cadre de campagnes. Il peut y avoir aussi combinaison des deux, des états sur papier facilitant les corrections manuelles.

4.1.1. Les corrections informatiques

Les corrections informatiques se veulent automatisées : il s'agit d'établir des programmes, des algorithmes de repérage et de correction des erreurs ou de dédoublonnage.

Les corrections par programmes sont à étudier par le Bureau et à mettre au point en collaboration avec les services informatiques, qui se chargeraient de l'exécution. De telles opérations ont déjà été programmées et exécutées par la cellule de maintenance actuelle de BN-Opale, avec succès. Il peut s'agir, à titre d'exemple de :

- mettre à jour les codes qui ont changé,
- corriger les zones codées (ex. : 008),
- corriger les erreurs de saisie repérables automatiquement (mettre le code de l'Autriche aux livres publiés à Vienne et qui sont en allemand, le code Venezuela aux notices en espagnol qui sont codées par erreur Vietnam, etc.),

- reformater les zones adresses et collation non différenciées (injection d'une 280 \$a devant chaque « In- » rencontré en zone 260),
- injecter un \$b dans les cotes des Imprimés qui possèdent dans le numéro séquentiel la mention « fol. » ou « p. »,
- injecter une barre de classement après la mention « Fol. » en début de titre (incunables),
- dédoublonner les notices APP qui comportent le numéro de la notice autorité.

4.1.2. Les corrections mixtes

Les services informatiques pourraient fournir des listings de notices présentant certaines caractéristiques repérables automatiquement. Cela faciliterait la mise au point de programmes de corrections automatiques quand elles sont concevables et/ou portent sur un grand nombre de notices, ou faciliterait le cas échéant le travail intellectuel avant corrections manuelles en ligne.

Par exemple pour faciliter les dédoublonnages :

- application à la base du logiciel « lot 6 » de JOUVE permettant de fournir des listings de « groupe de notices » ayant en commun au moins une cote commune³¹ ; l'exploitation manuelle de tels listings est le seul moyen fiable de procéder à des dédoublonnages pour des fichiers antérieurs à la normalisation,
- notices contenant une zone 983,
- titres identiques avec dates d'édition identiques.

Cependant le dédoublonnage par chaîne de caractère est trop risqué notamment pour les autorités personnes physiques homonymes.

Par exemple pour faciliter les corrections de données et de formatage :

- notices dont certaines zones normalement obligatoires sont vides ou inexistantes,
- notices ayant des données codées incompatibles avec les référentiels en vigueur,
- notices possédant des crochets carrés qui signalent un défaut de lisibilité ou l'absence d'information dans la notice d'origine mais sans doute pas sur l'ouvrage : [S. l. n. d. ?] ou [...].

Certaines actions peuvent relever d'actions informatiques programmées mais la plupart demandent non seulement l'oeil, mais la main, exercés, du professionnel du catalogage.

4.1.3. Les corrections manuelles

Nombre de problèmes et notamment les dédoublonnages ne peuvent être résolus que par des corrections manuelles, car toute destruction d'information doit être précédée de vérifications minutieuses ; de plus les données non présentes dans la notice conservée doivent y être copiées depuis les notices à détruire. D'aucunes peuvent y être présentes mais rédigées sous une autre forme (zone de notes, titre translittéré, etc). Aussi, des

³¹ Programme de dédoublonnage utilisé par la BnF et qui correspond à ce que recommandent les anglos-saxons et Y. DESRICHARD.

programmes informatiques ne peuvent suffire, une analyse intellectuelle reste irremplaçable et obligatoire pour une bibliothèque de ce type.

Concernant les corrections manuelles au long cours, il est possible d'énumérer les différents chantiers pour lesquels le matériel servant de base au travail existe. Il s'agit ainsi :

- de la reprise des lots dits de « retards »,
- du traitement des bordereaux concernant la CRI
- des bordereaux concernant BN-OPALE,
- du balayage systématique des index du cédérom de la phase 1³² pour corrections des plus grossières erreurs (cf. en annexe 5 des exemples d'écrans qui permettent un repérage facile d'erreurs),
- des corrections des erreurs contenues dans les échantillons qui ont servi aux contrôles qualité avant recettes des lots mensuels de saisie (pour chaque sous-lot le BQC possède les notices comportant des erreurs à corriger),
- des erreurs de code langue et de codes pays détectées lors du déchargement des données en vue de la publication du cédérom "Patrimoine imprimé des pays francophones du Sud",
- de l'exploitation des listings de rapprochement fournis par les prestataires pour rétablir les notices abusivement détruites et procéder aux dédoublonnages qui n'ont pu être exécutés du fait des spécifications. Rappelons que celles-ci ne pouvaient être que restrictives, l'ensemble du processus étant automatique pour les rapprochements Jouve préalables à l'édition commerciale du cédérom, pour les rapprochements entre marché 1 et 2 de la phase 2 avant la migration en 1997 et pour les rapprochements internes aux marchés 2, 3 et 4 de la phase 2.
- de l'exploitation des listes de notices contenant une zone 983 (qui indique la présence d'un doublon), listes fournies par le prestataire de contrôle (Bureau van Dijk) pour la 4ème partie de la phase 2.

Il est à noter que les diverses actions listées ci-dessus, quand elles sont menées manuellement, n'en font généralement qu'une, car l'essentiel est l'appréhension intellectuelle du catalogage, correct ou non d'un ouvrage donné.

En effet l'accès à une notice dont on veut corriger une cote entraîne la détection d'autres erreurs ou insuffisances dans cette notice, plus la détection de notices doublons, plus la constatation de l'incohérence des points d'accès avec les fichiers d'autorité. D'une simple correction de cote on passe à une amélioration de la notice et à un nettoyage du catalogue : en réalité tout se tient. La partie intellectuelle de ce travail est déjà faite pour un certain nombre de notices et archivée sous la forme des bordereaux de correction et des zones 983 des notices converties.

4.2. LES TRAVAUX SUPPLEMENTAIRES

Des compléments de conversion et de catalogage des UC (unités de conservation) orphelines sont également à prévoir. Il reste plusieurs fichiers et catalogues spécialisés qui n'étaient pas inclus dans le programme

³² Catalogue de la Bibliothèque nationale de France des origines à 1970. A son sujet on consultera : BEAUGENDRE, Anne-Marie. « Le Catalogue général des imprimés des origines à 1970 ».

de conversion rétrospective. Il faut vérifier fiche par fiche, notice par notice que leurs données figurent bien dans BN-Opale Plus afin d'assurer la complétude du grand catalogue. Selon l'ampleur des lacunes, les notices manquantes seront à créer en les cataloguant cette fois-ci, livre en main et en ligne, soit en envisageant des saisies complémentaires à l'extérieur.

4.2.1. Des compléments de conversion

Certains fichiers qui n'ont pas été retenus pour le chantier conversion rétrospective soit à cause de leur taille trop modeste, soit à cause de leur fort taux de doublonnage avec des catalogues convertis, soit parce qu'il s'agissait de fichiers internes de travail (qui ont été écartés par principe du chantier) doivent maintenant être examinés, confrontés avec le catalogue informatisé pour que les notices manquantes dans le SI soient saisies en interne. Il s'agit ainsi par exemple :

du Fichier des anonymes orientaux (ex-Hémicycle) : petit fichier des anonymes littéraires ou religieux dans les langues non-européennes.

du Fichier dit «des annuaires » (ex-Hémicycle) : fichiers contenant une majorité de publications à périodicité annuelle (almanachs, rapports annuels de collectivités, annuaires...) décrites surtout au fascicule et pour lesquelles il existe souvent, mais pas toujours, une description globale dans les fichiers de périodiques proprement dits (PR de la CRI ou BN-Opale).

de Fichiers de certains grands dons (Z. Maurois, Z Delaunay...)

de Fichiers internes de l'ex-Service des dons et actuellement stockés au D4, etc.

De même certains catalogues imprimés de fonds spéciaux devraient faire l'objet d'une saisie systématique à l'extérieur, ou en interne, selon leur taille. Leurs notices plus complètes devraient écraser les éventuelles notices déjà présentes dans le grand catalogue. Ces catalogues imprimés à saisir systématiquement sont par exemples :

le Catalogue des incunables

le Catalogue des règlements d'atelier

le Catalogue vietnamien 1960-1979

les Catalogues de la littérature africaine

les catalogues en alphabets non latins.

Il s'agit donc de combler les manques de la conversion. La base RECENSIO établie au début des travaux de migration pour recenser ces manques devrait permettre une recherche exhaustive des fichiers qui complètent, soit en totalité soit partiellement, les catalogues qui ont été convertis.

4.2.2. Le catalogage des UC

Enfin, et ce n'est pas le moindre des chantiers, toutes les UC non reliées et non reliables à une notice bibliographique, après recherches approfondies dans le catalogue, devraient être cataloguées. Il s'agit donc de créer les descriptions manquantes et d'améliorer les descriptions bibliographiques existantes pour permettre d'établir "l'épine dorsale"³³ dans

³³ Rappel : l'épine dorsale du SI ce sont les cotes, qui permettent de faire automatiquement un lien entre la description catalographique d'un ouvrage (notice

le SI pour tous les documents imprimés. Le but est double : d'une part de faire connaître l'ouvrage proprement dit, mal ou pas décrit du tout et d'autre part d'en rendre la communication automatisée possible via le SI. Cela revient à créer une épine dorsale pour des données de récolement qui n'ont pas leur équivalent dans la partie catalogue.

Le Bureau privilégierait le haut de l'épine dorsale se plaçant du point de vue "catalogue" et non du point de vue "récolement" puisque sa tâche consiste à améliorer le catalogue et non prioritairement à combler les lacunes et les erreurs du récolement. Outre cette tâche de catalogage, le Bureau devrait recataloguer certains ouvrages, notamment ceux dont la mauvaise qualité de la notice oblige à aller les voir en magasins ou à les faire venir dans le service.

4.2.3. L'enrichissement des notices de la base³⁴

Les notices de la conversion ne possèdent pour ainsi dire pas d'indexation matière (elles comptent tout de même environ 564 662 vedettes sujets non Rameau pour des fichiers réputés sans descripteurs matières !), 461 000 vedettes sujets Rameau et 38 000 indexations spécifiques à l'Histoire de France. C'est un inconvénient majeur surtout quand on sait que les chercheurs accèdent au catalogue surtout par ce mode³⁵. Une indexation rétrospective n'est pas envisageable car cela reviendrait presque à devoir recataloguer l'ensemble du fonds. A Richelieu, « à l'époque » des catalogues et fichiers papier, les usagers disposaient d'un fichier matière en Salle des catalogues. Ce fichier a été conservé et est toujours disponible en Salle de Référence Bibliographique. Mais l'idéal voudrait que ces vedettes matières figurent sur les notices informatisées. Quand les erreurs listées ci-dessus seront corrigées, le BQC pourrait enrichir les notices afin d'apporter une plus grande cohérence à l'ensemble. Mais, il est à craindre que cela ne soit jamais fait car cela demanderait de reprendre intégralement le fichier matière pour enrichir les notices. Or, ce fichier ne correspond plus aux normes d'indexation. Une possibilité aurait été de convertir ce fichier et procéder ensuite à des opérations automatiques de dédoublonnage. Cela avait été envisagé mais, après étude il était apparu que le fichier était par trop lacunaire. La « solution » la moins mauvaise consisterait à indexer les ouvrages qui feraient l'objet d'un recatalogage par le BQC pour cause de mauvaise qualité générale de la notice. Pour ce faire on peut envisager de récupérer -selon des modalités à étudier- l'indexation pratiquée sur le livre ancien par la British Library. Le reste des notices continuerait à être accessible par les accès principaux habituels.

De même, les notices de la rétroconversion ne disposent pratiquement pas de liens (62 000 liens avec des collections de BN-Opale). Aussi, afin de faciliter la consultation et notamment la production de la Bibliographie nationale française, il conviendrait de lier les notices bibliographiques aux

bibliographique) et l'enregistrement de son existence matérielle en magasin (données du récolement stockées autrefois dans la base SYCOMORE).

³⁴ On consultera : BOUDET, Isabelle, CLOAREC, Thierry. « Projet d'enrichissement pour le catalogue de la Bibliothèque de France »

³⁵ Cf. IHADJADENE, Majid. « L'accès sujet dans les catalogues en ligne : le cas des bibliothèques universitaires en France »

notices d'autorité chaque fois que le service serait amener à travailler sur une notice quelle qu'en soit le motif.

4.3. PROCEDURES DE TRAVAIL

Devant la masse des corrections à apporter, il paraît souhaitable de procéder comme ce qui se fait à l'inventaire rétrospectif ou à la réserve, c'est-à-dire de confier à chaque agent des tranches par lot à corriger. Cela devrait se faire sur la base du volontariat tout en émettant la réserve qu'il serait préférable que les agents ayant supervisé et suivi la saisie d'un lot s'occupent également de ses corrections. La spécialisation pour la préparation des lots et le travail courant lors de leur saisie ont permis de les connaître. Il serait regrettable de ne pas tirer profit de cette connaissance acquise sur le tas. Les principes de rédaction des catalogues et des fichiers se sont perdus aussi il ne faudrait pas dilapider cette connaissance acquise au cours de ces 10 années de conversion. Cela permettra à chacun une responsabilisation plus motivante sur un lot et surtout de pouvoir voir le travail avancer.

Au sein du Bureau, le temps de travail semble devoir être obligatoirement réparti en 3 types de tâches distinctes, selon des modalités définies périodiquement par le responsable du service :

- les corrections d'erreurs signalées par le personnel de la BnF ou par des lecteurs (ex. : messages à propos du catalogue sur Internet),
- le programme personnel de corrections selon certains critères ou programme personnel de tri et de saisie complémentaire de fichiers/catalogues non convertis,
- le service public en banque de salle de lecture Rez ou Haut-de-Jardin (2 séances hebdomadaires au maximum).

Les tâches spécifiques au service culminent par la saisie de données catalographiques correctes en format InterMarc intégré et selon les normes et consignes applicables à la BnF. La saisie est précédée de recherches approfondies dans le catalogue BN-Opale Plus mais aussi dans les anciens catalogues de l'établissement (bases anciennes archivées, cédéroms BnF, catalogues imprimés, fichiers), dans les ouvrages de références du service et des services voisins (Inventaire rétrospectif, Réserve, Département de la recherche bibliographique...) et dans les fonds de la BnF (magasins et libre-accès).

Ci-après on trouvera un tableau récapitulatif des tâches à effectuer.

PANORAMA DES TACHES

| TRAVAIL COURANT | | TRAVAIL PROGRAMME | |
|--|--|--|---|
| Gestion et Coordination | Corrections 1830-1999 | Manuel | Corrections automatiques |
| Réception, tri, enregistrement et répartition des bordereaux de correction et messages reçus sur le web concernant la qualité du catalogue | Tri, enregistrement, distribution des demandes | A partir de documents papier préexistants : Bordereaux de corrections/questions sur le catalogue listings de contrôle qualité de la CRI listings des rapprochements CRI anomalies non traitées CRI (100 000) | <u>Reformatages</u> (Ex. : zone 260 des notices SU de la CRI) |

| | | | |
|--|--|---|--------------------------------------|
| | | - liste des zones 983 dans la CRI phase 2 marché 4 (doublons repérés) | |
| Établissement et diffusion de statistiques mensuelles et annuelles de la collecte des demandes | Traitement : recherches et corrections effectives dans BN-OPALE PLUS | A partir des listings d'anomalies des diverses migrations : de BN-OPALE - de la CRI | Données (Ex. : codes pays et langue) |
| | | Nettoyages à partir des index (titres, cotes) | |
| Statistiques mensuelles et annuelles du traitement des corrections courantes et systématiques (par BQC et autres services) | Retour aux émetteurs avec explications et références (délai maximum visé : la semaine) | Tri des fichiers et catalogues imprimés non inclus dans la CRI (Ex. : catalogue vietnamiens, Anonymes orientaux, fichier des "annuaires ex-hémicycle) | |
| | | Manuel, d'après listings produits par DSI | |
| | | Repérage des problèmes potentiels (Ex. : titres et dates identiques) | |

Afin de lancer l'opération de corrections, le Bureau doit créer un :

- tableau de recensement des corrections à effectuer sur les données de la CRI,
- programme des opérations à effectuer (annuel et/ou trimestriel),
- formulaire de demande de corrections de données,
- formulaire de demandes de renseignements sur les données adressées à des services hors Tolbiac.

A partir de tous ces éléments, il est possible d'établir un tableau des fonctions du Bureau Qualité du Catalogue³⁶.

5. LES AGENTS DU BUREAU ET LEURS ATTRIBUTIONS

5.1. LES COMPETENCES REQUISES

Le maintien de la cohérence de la base n'est pas un simple travail de saisie mais il nécessite aussi un jugement, des aptitudes techniques et le recours aux ouvrages, des recherches en magasin, des recherches dans les anciens catalogues et les images des anciennes bases (cédéroms) seront nécessaires dans de très nombreux cas. Ce travail est beaucoup plus complexe qu'il apparaît à première vue et peut être même plus complexe que le catalogage courant.

Afin de mener des programmes de corrections au long court tout en assurant les demandes quotidiennes, il semble donc nécessaire pour la qualité du travail de posséder certains prérequis comme :

- d'avoir une expérience de l'informatique documentaire et notamment du catalogage en ligne,
- d'avoir un intérêt pour les catalogues et les bases bibliographiques informatisées,
- de maîtriser les normes et les formats de catalogage notamment l'Intermarc et l'Intermarc intégré,

- de posséder une connaissance approfondie des principaux fichiers et catalogues qui constituent le grand catalogue du SI,
- de connaître les fonds anciens de la BN et les modes de classement de la BnF (cotation Clément, cotes inventaires et lettres cataloguées)³⁷,
- d'avoir une grande pratique des difficultés de catalogage et du catalogage informatisé,
- d'être rigoureux et méthodique,
- d'avoir le sens du travail en équipe et le goût des relations avec les autres services et le public.

5.2. L'IMPLICATION EN PERSONNEL

Le nombre de corrections à mener et la demande accrue de maintenance prévue implique du personnel dédié à cette tâche. Il faut donc que les effectifs du Bureau Qualité du Catalogue soient non seulement permanents, mais importants et qualifiés.

Le chantier de conversion rétrospective s'étant achevé, le personnel de l'ex-SCCR se trouvait être disponible³⁸. Son affectation au nouveau service de cohérence du catalogue, le Bureau Qualité du Catalogue, s'est avéré être un bien car l'expérience a montré que la Cellule de Cohérence de la base (BN-Opale) avait du mal à assurer sa tâche faute d'effectifs suffisants et stables³⁹, et d'une proportion adéquate de professionnels qualifiés.

La Cellule de Cohérence de la base, s'est ainsi vue renforcer d'une équipe de 7 contractuels de catégorie B et A et d'un encadrement composé d'un conservateur général, un bibliothécaire et un contractuel catégorie A. Le Bureau dispose ainsi d'une équipe de 13 personnes, ce qui semble pour lors, dans cette période de restrictions en personnel et budgétaire, suffisant pour les opérations à mener sachant que rares sont les bibliothèques pouvant se permettre de posséder un tel service. Le travail étant très technique et peu gratifiant, le Bureau déplore l'absence de mutation interne à son profit⁴⁰. Cependant ce nombre d'agents semble satisfaisant eu égard le chiffre des agents affectés à ce genre de tâche dans d'autres établissements. A titre d'exemple, nous pouvons citer le service « Coordination bibliothéconomique et informatique » de la Bibliothèque Publique d'Information qui compte pour une base bien moindre, 10 agents affectés exclusivement à ce type de tâche.

³⁶ Cf. annexe 4.

³⁷ Rappelons que la nouvelle cotation n'a commencé à être appliquée qu'en 1997. Sur l'histoire de la cotation à la BnF on consultera : BLASSELLE, Bruno. « Classement des savoirs et classement des livres de la Bibliothèque nationale de France ». In *Tous les savoirs du monde : encyclopédies et bibliothèques, de Sumer au XXIe siècle* / sous la dir. de Roland Schaer. - Paris : Bibliothèque nationale de France : Flammarion, 1996, p. 194-199 et PORTES, Laurent. « L'ordre des livres selon Nicolas Clément ». In *Mélanges autour de l'histoire des livres imprimés et périodiques*. - Paris : Ed. de la Bibliothèque nationale de France, 1999, p. 64-92

³⁸ L'encadrement du Bureau, grâce à l'expérience acquise en matière de conversion de catalogues de bibliothèque pendant cette opération de dix années assurera une expertise, au service de l'ensemble de l'établissement voire d'établissements extérieurs (aide fournie pour la rédaction des cahiers des charges de l'Arsenal, des Estampes et de la Bibliothèque centrale des Musées nationaux). Ces contacts avec l'extérieur lui permettront d'enrichir la pratique du bureau par celle des autres.

³⁹ L'équipe ne comptait qu'un titulaire pour 4 contractuels.

⁴⁰ A part une BA devant arriver en janvier 2000.

5.3. LES FORMATIONS A METTRE EN ŒUVRE

A moins de mutations internes pour d'autres services de l'établissement, de réussites de concours ou de démissions au profit d'autres établissements⁴¹, le BQC a la chance d'être composé des agents de la conversion rétrospective, c'est-à-dire des agents qui connaissent le mieux le travail à effectuer et les fichiers à traiter.

Le BQC du fait de la réaffectation des agents issus du SCCR bénéficie d'un personnel compétent, formé et motivé. Il se trouve être immédiatement opérationnel en tout cas pour ce qui relève du travail à effectuer sur les fichiers saisis.

La seule formation complémentaire nécessaire est la formation commune à tous les futurs catalogueurs dans le SI, ce qui relève des applications de la V2. Il s'agit de permettre au personnel de cataloguer et d'indexer dans la base et donc de lui donner les accès et attributions au moins des catalogueurs du Dépôt légal sinon des bureaux autorités.

De plus ce personnel est motivé pour accomplir un type de travail qui lui est familier car il s'apparente au traitement des anomalies de la conversion, et il a acquis de réelles qualifications dans la connaissance des fonds patrimoniaux et de leurs anciens catalogues. Cette motivation pour ces travaux qui peuvent paraître ingrats est sous-tendue par un souci du public, qui s'est traduit par le fort taux de volontariat pour le service public.

Le travail dans les banques de salle l'a rendu d'autant plus conscient des problèmes de communication suscités par les erreurs de cotes notamment, et a montré combien l'expérience acquise au sein du SCCR était un atout très important, utile pour résoudre les problèmes : sachant mieux analyser l'origine des dysfonctionnements, le service rendu au public s'est trouvé plus efficace.

5.4. LES RESPONSABLES DES TRAVAUX DE CORRECTION

Comme on l'a vu, le service bénéficie du personnel qui l'a vu naître ou qui l'a rejoint lors de son développement. La connaissance approfondie que les agents du Bureau ont acquis de la majeure partie des fichiers et catalogues qui constituent le grand catalogue du SI par le chantier de conversion, sa grande pratique des difficultés de catalogage et du catalogage informatisé, des fonds anciens et des cotes BN est un atout très important pour la qualité du travail.

Aussi il est irréaliste d'attribuer les corrections à ceux qui en sont les auteurs, par exemple aux divers services de catalogage courant quand ils en commettent, puisqu'ils sont déjà très chargés, et sont ponctionnés par les lourdes nécessités du service public : l'expérience montre que le travail courant l'emporte toujours sur le travail rétrospectif. Il n'est pas non plus opportun de le faire à cause de la méconnaissance d'une partie croissante du personnel des fonds patrimoniaux anciens et de leur système de cotation, car l'évolution de l'établissement fait que beaucoup n'ont connu que le site Tolbiac sans avoir jamais fréquenté les magasins de Richelieu, source irremplaçable de connaissance des cotes. Rappelons que les fonds relevant de

⁴¹ Rappelons que le service n'est composé que de contractuels à part son chef, conservateur général et de son adjoint, bibliothécaire stagiaire (moi-même).

la cotation Clément constituent l'écrasante majorité du catalogue dans le SI, car la totalité des cotes de la conversion rétrospective des imprimés et la grande majorité des cotes de BN-Opale en relèvent.

Par conséquent la logique veut que ce soit le service qui en a les missions et le personnel, qui s'occupe de la cohérence de la base et des corrections. Cela dit, il est nécessaire de maintenir, à côté du Bureau Qualité du Catalogue pour la correction des données migrées et à côté des services gérant les divers fichiers d'autorité, un service veillant à la cohérence des notices bibliographiques créées en ligne dans le SI, c'est-à-dire des notices n'ayant pas eu à subir des opérations de conversion ou de migration. Il s'agit là plutôt d'une fonction de suivi de formation et de veille, avec délégation de l'exécution des corrections aux producteurs de notices eux-mêmes, dans un but pédagogique. Réunis au sein du Service de Coordination Bibliographique, la cloison ne devrait pas être étanche entre les deux services.

Si les corrections incombent au Bureau, cela n'empêche pas les collaborations voire les répartitions du travail.

5.5. LES COLLABORATEURS DU BQC

Les partenaires fonctionnels à envisager pour effectuer des contrôles et des corrections dans la base sont de deux types : ceux avec lesquels il faudra se répartir la tâche et ceux qui interviendront comme collaborateurs du Bureau pour effectuer sa tâche.

Plusieurs services de la BnF sont susceptibles de pouvoir et devoir corriger des notices dans la base, en plus du Bureau Qualité du Catalogue. En effet, le travail de recherche et de catalogage des notices bibliographiques susceptibles de correspondre aux UC (unités de conservation) restées isolées après constitution de l'épine dorsale est à partager entre la Direction des Collections et le Bureau. Comme il existe deux normes AFNOR pour cataloguer les monographies, il est clair qu'il faut des compétences différentes, au moins en partie, pour cataloguer les livres selon leur époque de parution. A la BnF les compétences pour les incunables se trouvent à la Réserve ainsi que les compétences pour les livres rares et remarquables (qui ont justement, ou auront prochainement, une cote "Rés."), les compétences pour cataloguer les livres anciens sont regroupées à l'Inventaire rétrospectif. D'autre part, les agents du Bureau ne possèdent pas la qualification pour manipuler les cotes, aucun n'émanant par exemple du chantier récolement. Aussi, selon la nature de l'UC (recueil, monographie « moderne », etc.), le partage de la responsabilité des corrections est à mettre au point avec :

- les différents bureaux responsables des autorités au sein du SCB dont le degré de qualification et de compétence ne peut se substituer à celui du BQC,
- la Réserve des livres rares, pour le traitement des incunables, les ouvrages antérieurs à 1500,
- le Service de l'inventaire rétrospectif (ex - Inventaire général) qui a en charge le catalogage du livre ancien (1500 - 1811⁴²) qui exige une telle connaissance du livre ancien qu'il paraît peu pertinent de le confier au BQC

⁴² Il s'agit des dates de sa tranche de traitement mais il est envisager de l'élargir éventuellement jusqu'en 1830 à l'image de la tranche chronologique habituelle de la Réserve qui a été modifiée pour son chantier d'accroissement.

qui bien qu'ayant converti la majorité des fonds anciens et ayant par là eu une approche non négligeable du livre ancien, ne peut se substituer aux agents ayant bénéficié de formations spécifiques à son traitement. Une connaissance empirique ne peut remplacer l'expérience acquise par ce service,

- le service des périodiques pour les notices relevant de ce type de support sachant que le fonds des imprimés en contient beaucoup : les périodiques classés in-folio, etc.,
- le service des langues orientales pour toutes les notices en caractères non latins,
- le Département de l'audiovisuel,
- les services ayant pris la suite du service de récolement pour la correction des UC issues de Sycomore,
- des prestataires extérieurs pour des saisies complémentaires de catalogues et fichiers.

Le professionnalisme reconnu des services en cause devrait les inciter à accepter ces responsabilités et les devoirs qui s'y rattacheraient : c'est-à-dire le devoir de corriger certaines notices, en dehors et en plus des programmes de recatalogage. Le partage des responsabilités et des tâches pourrait se faire selon un principe chronologique : il aurait le mérite d'être strict, clair et de correspondre aux attributions déjà en œuvre pour ce qui est du recatalogage en cours et selon le principe du type de document. Ce partage des tâches avec la DCO devrait faire l'objet d'un contrat détaillé et écrit, une sorte de « charte de travail »⁴³.

Pour ce qui est des collaborateurs du BQC qui auront à travailler pour ou avec lui, il faut citer le Département des Systèmes d'Information (de la DSR) pour mener des programmes informatiques de nettoyage, de dédoublonnage, fournir des listings permettant un travail systématique, etc.

L'opération relevant des missions de l'établissement et devant pour l'ensemble des procédures se dérouler en interne, il n'y a pas de partenaires financiers à envisager. D'autant que les budgets spéciaux attribués dans le cadre des chantiers de l'EPBdF sont clos. Seules quelques saisies supplémentaires de conversion pourraient faire l'objet de marchés avec des prestataires extérieurs. N'étant pas envisagés, il est pour lors impossible de savoir s'ils seraient effectués sur le budget de la direction, de l'établissement ou feraient l'objet d'une demande spécifique aux tutelles.

5.6. L'ECHEANCIER DES TRAVAUX

Les travaux de correction à apporter dans la base sont des travaux au long cours à assurer dans le SI quand les données seront accessibles en ligne, et non de chantiers transitoires à effectuer au moment de divers versements et migrations. Les travaux de conversion rétrospective s'étant achevés à l'été 1999, le BQC aurait pu entamer ses travaux dès l'automne. Mais le marché de construction du SI et notamment de la mise en oeuvre de la V2, c'est-à-dire la disponibilité du module P.C.A. (Production du CAtalogue) du SI qui aurait du permettre la création et la modification en ligne de notices a été résilié (le 23 juillet 1999). Ce n'est donc, en l'état actuel du calendrier du SI, qu'à partir de l'automne 2001 que les corrections

⁴³ comme le dit GUY, Michèle. *Rapport annuel APP : 1998* [: document interne]

manuelles au long cours pourront commencer. Ce report de livraison va permettre de mener des travaux préparatoires, de former l'équipe à ses futures tâches, de se remettre à niveau en ce qui concerne les normes de catalogage.

Le volume du travail à prévoir fait écarter d'emblée une structure provisoire qui pourrait mener à bien une tâche bien balisée dans un calendrier défini. Même quand les performances du module de catalogage dans le SI seront connues pour pouvoir évaluer le temps moyen nécessaire en ligne à la correction d'une notice ou à la suppression d'un doublon, il ne sera pas vraiment possible de chiffrer tous les travaux de correction en homme/mois. Notons aussi que l'accès aux ouvrages et aux fichiers d'origine est particulièrement lourd sur le site Tolbiac (dispersion dans de nombreux magasins, procédures d'accès plus contrôlées) ce qui peut ralentir la tâche.

Cependant certains dysfonctionnements semblent à effectuer d'urgence. Les priorités pour les données bibliographiques provenant de la conversion sont :

- d'assurer ou d'améliorer l'accès aux notices des ouvrages qui étaient repérables dans les catalogues papier d'origine et donc de restaurer cet accès s'il a été compromis ou détruit par la conversion ou s'il est affecté par des erreurs qui les rendent inatteignables lors d'une recherche,

- de rendre efficace la communication des ouvrages déjà décrits dans le catalogue et donc susceptibles d'être demandés par des lecteurs par les nouvelles procédures du SI et donc de constituer manuellement les "épines dorsales" qui n'ont pu l'être à la migration: c'est donc par les corrections de cotes que l'on doit commencer,

- d'améliorer la recherche bibliographique et documentaire en créant les accès qui n'existaient pas dans les catalogues rédigés avant l'existence des normes de catalogage.

Pour ce faire il est possible d'établir un ordre de priorité des activités à mettre en place. Certains points sont à aplanir avant d'entamer le travail. Ainsi, il convient de :

- former l'équipe au catalogage dans BN-Opale Plus (Intermarc intégré, PCA : production du catalogue, normes, fichiers d'autorité),

- de délimiter le champ d'action du Bureau et sa collaboration avec les autres services travaillant au catalogue,

- de préparer les corrections manuelles déjà repérées de sorte à en permettre une saisie efficace dès la livraison de la V2 (préparation intellectuelle et matérielle). Il faut recenser les campagnes systématiques de corrections manuelles à faire sur les données de la CRI (ce qui est déjà fait pour les données en provenance de BN-Opale ; ex. : oubli de la barre de classement derrière les articles initiaux, erreurs de codes pays) et recenser les documents papier, autres que les bordereaux de corrections, permettant de repérer des erreurs ou des manques (listings de contrôle qualité de la CRI, liste des cotes à lettrage grec Pi avec les cotes postérieurement attribuées mais non reportées dans les catalogues, éditions des notices ayant subi des programmes de rapprochements en vue de dédoublonnages chez les prestataires de saisie, cotes non normalisables lors des migrations dans le SI),

- de mettre en place un circuit efficace pour les erreurs détectées au jour le jour. Au sein du Bureau, les procédures de travail sont à mettre en

place, du moins les principes de départ en l'absence de l'outil de travail ou d'une maquette consultable sur une plate-forme provisoire permettant de définir des procédures précises d'intervention sur le catalogue,

- de mener des réflexions sur les travaux au long cours (corrections par programmes automatiques, vérifications et saisies complémentaires...). Le Bureau peut d'ores et déjà recenser les fichiers et catalogues imprimés non pris en compte dans le chantier et susceptibles de contenir des notices uniques afin de combler les manques dans BN-Opale Plus (ex. : Fichier des Almanachs de l'ancien Hémicycle de Richelieu, Catalogues imprimés du fonds vietnamien), préparer une méthode de tri à rendre effective dès que tous les chargements de la conversion rétrospective des imprimés auront été effectués et une méthode de saisie (copie des fiches, recatalogage livre en main...). Le plus important semble d'étudier les possibilités d'intervention permises par le SI afin de peser sur les cahiers des charges à venir pour la passation de nouveaux marchés et notamment la prise en compte d'outils d'identification tels le SQL,

- et d'intégrer l'équipe à la marche quotidienne de l'établissement par une participation généralisée au service public en salles de lecture.

En ce qui concerne l'encadrement, il doit assurer le suivi de la migration pour les chargements à valider, et les jeux d'essai à constituer.

Pour ce qui est du calendrier exact des corrections, d'un échéancier, il est difficile d'établir un ordre de succession précis des travaux tant les divers travaux listés ci-dessus s'interpénètrent entre eux. De même il est délicat de fixer un calendrier des dédoublonnages à mener, puisque c'est au cours des recherches que l'on est amené à constater les défauts de tous ordres du catalogue. Les incertitudes du calendrier (des formations dispensées par la BnF, de la disponibilité de PCA...) rendent la programmation précise encore aléatoire.

Des corrections seront à porter immédiatement. Il s'agit de celles qui émanent d'usagers du catalogue. La collecte des demandes de correction est organisée sous forme de formulaire que la personne qui a détecté une erreur doit renseigner et faire parvenir au BQC. Le formulaire de demande devrait bien entendu être disponible sur Lotus Notes (et sur le site web de la BnF ?) ; mais la version papier semble indispensable pour les lecteurs et le personnel ne disposant pas encore de Lotus Notes ou d'Internet.

La rapidité est un service au public (et au personnel BnF, le catalogue étant un élément majeur du SI, instrument de travail commun) qu'il faudrait considérer comme une priorité. Le service pourrait s'engager sur les délais (48 heures par exemple) et retourner l'information sur la suite donnée au demandeur, exactement comme on le fait pour les demandes d'acquisitions. Les délais de correction doivent être courts pour encourager les utilisateurs et notamment le personnel à faire part de ses remarques, à transmettre ses fiches de correction.

6. DES REFERENCES EXTERIEURES : EXEMPLES DE LA BDIC ET DE LA BPI

Afin de me renseigner sur les pratiques d'autres établissements concernant le sujet de ce PPP, il m'a semblé intéressant et profitable d'étudier un service assurant le genre de maintenance que projette le Bureau Qualité du

Catalogue. De ce fait je me suis rendu en stage à la Bibliothèque de Documentation et d'Information Contemporaine (BDIC) et en visite à la Bibliothèque Publique d'Information (BPI) du Centre d'Art Georges Pompidou.

6.1. ANALYSE DE CES REFERENCES

Plutôt que de décrire les deux établissements et leurs pratiques, il me paraît préférable d'en relever les constantes et les divergences pour en tirer des enseignements.

6.1.1. Les convergences

La Bibliothèque de Documentation Internationale Contemporaine et la Bibliothèque Publique d'Information possèdent toutes deux :

- un système GEAC : Advance à la BDIC et Gliss 9000 à la BPI,
- un service dédié au contrôle : « Coordination bibliothéconomique et informatique » à la BPI et « Service informatique et réseaux » à la BDIC,
- un fonds multimédia et multisupports,
- un contrôle visant à servir le public et donc centré sur les accès, les autorités,
- des codes d'accès pour travailler dans la base en fonction des aptitudes,
- des contrôles automatiques de la syntaxe et des doublons,
- une procédure manuelle de correction des autorités. La complexité des corrections à apporter en terme intellectuel interdit la rédaction de programmes informatiques de correction qui seraient trop complexes et pas assez rentables,
- des types de problèmes à résoudre bien identifiés (autorités dérivées d'OCLC à traduire en français à la BDIC, uniformisation des autorités issues de Laval avec celles de Rameau à la BPI).

6.1.2. Les divergences

Cependant des divergences existent entre les deux bibliothèques :

- le nombre d'agents pour le service. A la BDIC, le contrôle qualité de la base est effectué par des acquéreurs-catalogueurs qui partagent leur temps entre leurs différentes activités alors qu'à la BPI le service compte 10 personnes s'occupant que du contrôle,
- l'identification des formes doublons est automatique à la BDIC : la base permet d'éditer un listing d'autorités qui peuvent être fusionnées,
- un système plus souple à la BDIC : fusions automatiques d'autorités possibles ainsi que travail dans la partie autorité de la base et non dans la partie bibliographique comme à la BPI,
- la BPI ne possède pas de formulaire de correction,
- la BPI ne contrôle pas le catalogue courant,
- le catalogue de la BDIC est composé de différentes strates de catalogage (dérivation de BN-Opale ; dérivation d'OCLC ; rétroconversion ; saisie).

6.2. LES ENSEIGNEMENTS A PARTIR DES POINTS FORTS ET FAIBLES

Le principal élément positif du travail de la BDIC est la procédure semi automatisée permise par le logiciel. Cependant la connaissance nécessaire de la base pour effectuer des corrections complexes exige des formations sur les autorités-matières RAMEAU pour les résoudre car sinon, il faut multiplier les requêtes pour accéder à la bonne forme, faute de bien connaître le langage d'indexation alors que la spécialisation de la BPI permet de soulager les catalogueurs. A la BDIC le dédoublonnage est long car il est impossible d'effectuer une modification en plus d'une fusion alors même que le système rend obligatoire la correction complète de la notice pour pouvoir enregistrer la fusion. Aussi il faut renouveler les procédures d'interrogation, d'affichage de la notice puis de correction... Cela est long et fastidieux car le système multiplie les commandes pour verrouiller les opérations, les complexifie pour les sécuriser.

Pour ce qui est de la BPI le principal point fort est le service spécialement dédié au contrôles même si l'on peut regretter l'absence de contrôle du catalogage courant -qui se justifie pour cette bibliothèque au desherbage fréquent- de même que l'absence de formulaire de demande de correction.

Ces différents éléments permettent de tirer de rares enseignements pour le Bureau Qualité du Catalogue tant l'ensemble des procédures étudiées avaient été prévues par le service. Les points convergents non recensés devraient être appliqués. Sinon l'absence de procédure automatique d'identification pour les deux établissements laisse à penser que le travail à partir des sources déjà répertoriées peut suffire et qu'il ne faut pas trop regretter l'absence d'un programme « miracle » permettant de repérer toutes les fautes inconnues.

7. LE SUIVI DE LA COHERENCE DES DONNEES

7.1. LES OUTILS D'EVALUATION A METTRE EN PLACE

Le Bureau Qualité du Catalogue doit mettre en place des outils d'évaluation de son activité. La satisfaction des besoins des utilisateurs et notamment celle de leurs recherches dans la base est un point important du contrôle de la qualité. C'est pourquoi, des évaluations se doivent d'être statistiques, quantitatives et qualitatives (sur la satisfaction sur les temps de réponse du service par exemple).

Evidemment, afin de suivre l'activité, il convient d'établir des statistiques mensuelles et annuelles de la collecte des demandes, du traitement des corrections courantes et systématiques effectuées aussi bien par le BQC que par les autres services en cause ainsi que des réponses fournies. Le suivi de l'activité peut être assisté par l'outil informatique.

7.2. L'OBJET DE CES BILANS ET LEURS DESTINATAIRES

Ces outils d'évaluation sont destinés en premier lieu au service lui-même afin d'évaluer son activité et d'éviter une démotivation des agents du BQC potentielle avec la visibilité tardive du travail. En second lieu ils sont destinés à la hiérarchie afin de la tenir informée de l'état d'avancement des travaux et justifier par là son action tant le travail de maints services de

la BnF ne lui semble pas concret tant qu'il n'est pas chiffré d'autant que la masse des améliorations à apporter rendra le travail du bureau invisible pendant longtemps. Enfin, il convient d'en faire part à tous les utilisateurs car ces travaux n'atteindront leur objectif que s'ils sont diffusés aux usagers, si l'information est largement relayée.

CONCLUSION

« Tous les fonds imprimés antérieurs à 1970, qui étaient jusqu'ici répertoriés de façon hétérogène, couvrant des tranches chronologiques plus ou moins larges et parfois même non consultables par le public, se retrouvent désormais dans un catalogue unique et interrogeable à distance, offrant à la communauté scientifique et professionnelle un outil indispensable à la recherche ainsi qu'un réservoir bibliographique de 7 millions de références à toutes les bibliothèques de France et du monde. Un bel exploit, mais qui aura besoin d'un bel outil informatique pour fonctionner »⁴⁴.

A la diversité des catalogues et fichiers de la BN de Richelieu dont la complexité a fait naître des manuels et des guides à l'usage du lecteur, s'est substituée une base unique de données, BN-Opale Plus qui offre un accès unifié à l'ensemble des fonds imprimés et audiovisuels de l'établissement. Mais si le support est unique, le contenu est variable dans sa qualité de par l'hétérogénéité des sources qui le compose et son mode d'alimentation essentiellement du à la rétroconversion des fichiers et catalogues. Aussi, comme le laisse entendre Laurence Santantonios, un travail visant à améliorer la cohérence intellectuelle et à enrichir le catalogue est à mener afin de faire du catalogue un bel outil. Il convient donc d'abord d'identifier les erreurs puis de les corriger. Car en effet, comme le dit Gabriel Naudé « qu'il faille apres tant de peine & de defpenfe cacher toutes ces lumieres fous le boiffeau, & condamner tant de braues efprits à vn perpetuel filence & folitude, c'eft mal recognoiftre le but d'vne Bibliotheque.»⁴⁵.

Aussi, pour le Bureau Qualité du Catalogue, acteur majeur de cette politique d'amélioration, comme pour l'ensemble des utilisateurs du catalogue, BN-Opale Plus n'en est qu'à ses débuts.

⁴⁴ SANTANTONIOS, Laurence. « Embellie de l'été à la BnF »

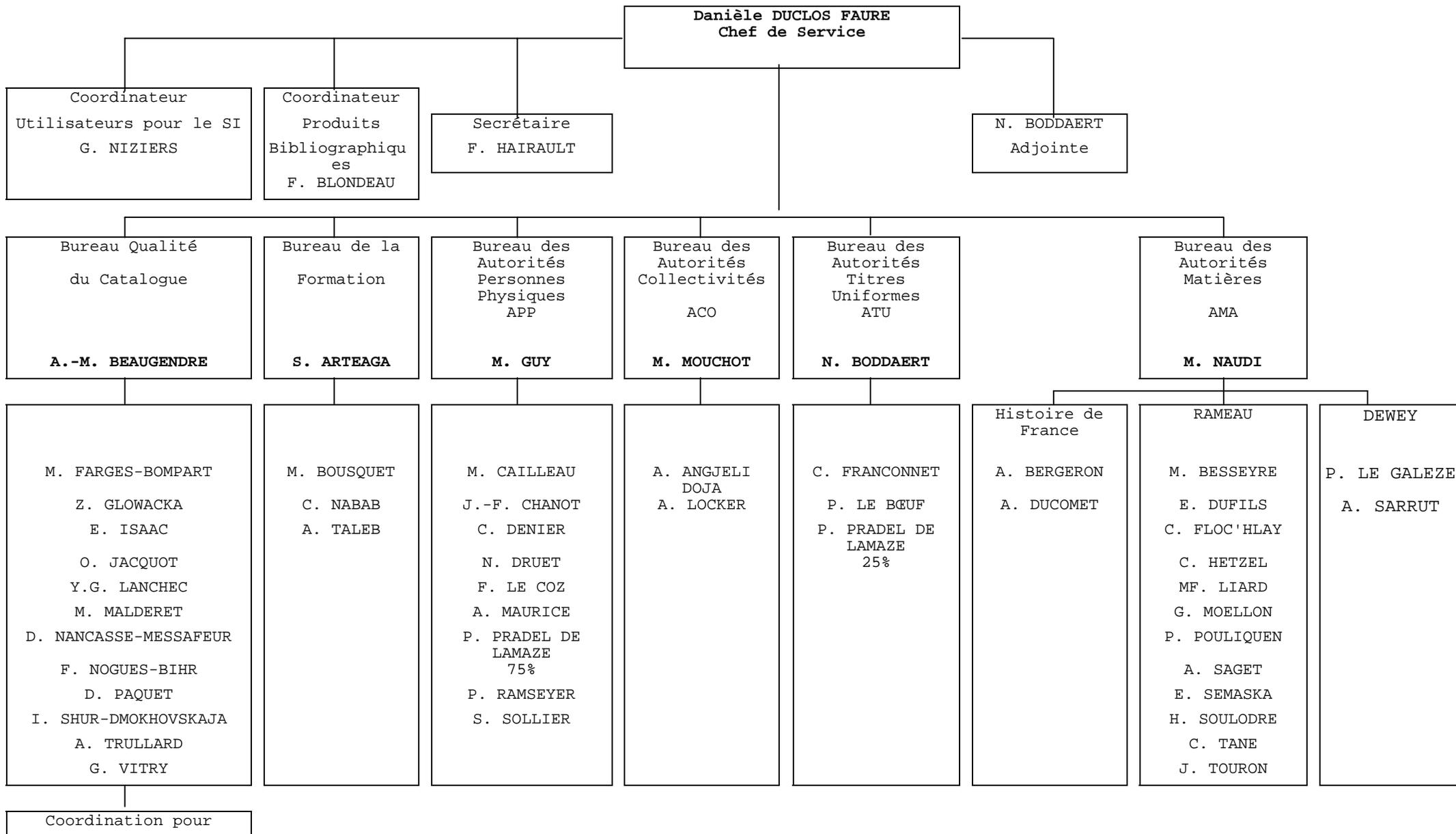
⁴⁵ NAUDÉ, Gabriel. *Advis pour dresser une bibliothèque*

ANNEXES

1. Organigramme du SCB
2. Volumétrie maximum d'erreurs calculée d'après le NQA
3. Cotes en erreur
4. Bureau qualité du catalogue (BQC) : panorama des fonctions
5. Exemples d'index du cédérom *Catalogue général des imprimés des origines à 1970*
6. Exemple de grappe issue du SI : Balzac

1. Organigramme du SCB

DIRECTION DES SERVICES ET DES RESEAUX
 AGENCE BIBLIOGRAPHIQUE NATIONALE - SERVICE DE COORDINATION BIBLIOGRAPHIQUE



Les BU
N. DADMEHR
T. FARAGASSO

2. Volumétrie maximum d'erreurs calculée d'après le NQA

TABLEAU 1
VOLUMETRIE MAXIMUM calculée d'après le NQA

Pour des raisons d'hétérogénéité des données qui servent de base aux calculs ceux-ci ne sont que des approximations, néanmoins significatives de l'importance du travail à engager. On s'est efforcé quand c'était possible de ne traiter que des notices de type bibliographique, excluant les notices d'autorité, considérant que le travail sur ces notices était du ressort des services spécialisés du SCB. On a utilisé dans l'ordre décroissant de priorité :

- les chiffres provenant des statistiques de la migration pour V1
- les chiffres de la saisie après dédoublemnages internes à la CRI pré-migration V1.*
- les chiffres d'estimation de la saisie à faire ou en cours: **

| Nom du catalogue | Type d'erreur | NQA = nombre d'erreurs maximum tolérées pour un nombre de notices | Nombre total de notices bibliographiques dans le SI ou de zones | Nombre maximum d'erreurs ou de notices avec des erreurs |
|------------------|--|---|---|---|
| CG | Catalogue général Auteurs (232 vol.) | | | |
| | cotes | 2 sur 1 000 | 1 936 097 cotes | 3 872 cotes |
| | vedettes auteurs et titres factices | 4 sur 1 000 | 1 926 771 zones 100/700/141/143 | 7 708 vedettes |
| | formatage = zonage | 1 sur 1 000 | 1 776 878 notices | 1 777 notices |
| | frappe, codes langue et pays, barre de classement des titres | 60 sur 1 000 | 1 776 878 notices | 106 613 notices |
| | total CG hors cotes | | | 116 098 |
| | total CG | | | 119 970 |
| SU | Supplément sur fiches au CG Auteurs | | | |
| | cotes | 2 sur 1 000 | 1 018 205 cotes | 2 036 cotes |
| | vedettes auteurs et titres factices | 8 sur 1 000 | 1 266 515 zones 100/141/143/700/710 | 10 132 zones |
| | formatage = zonage | 2 sur 1 000 | 1 001 441 notices | 2 002 notices |
| | frappe, codes, barres | 120 sur 1 000 | 1 017 297 notices | 122 076 notices |
| | total SU hors cotes | | | 134 210 |
| | total SU | | | 136 046 |
| LX | Catalogue 1960-1969 | | | |
| | Cotes | 2 sur 1 000 | 330 155 cotes | 660 cotes |
| | vedettes auteurs et titres factices | 6 sur 1 000 | 418 046 zones 100/110/141/143/700/710 | 2 508 zones |
| | formatage = zonage | 2 sur 1000 | 337 584 notices | 675 notices |
| | frappe, codes, barres | 90 sur 1000 | 337 584 notices | 30 383 notices |
| | total LX hors cotes | | | 33 566 |
| | total LX | | | 34 226 |
| PR | Périodiques avant 1960 | | | |
| | cotes | 2 sur 1 000 | 225 420 cotes | 451 cotes |
| | titres | 4 sur 1 000 | 216 615 notices bibliographiques (J et H)t | 867 notices |
| | formatage = zonage | 1 sur 1 000 | 216 615 notices | 217 notices |
| | frappe, codes, barre de classement du titre | 60 sur 1 000 | 216 615 notices | 1 300 notices |
| | total PR hors cotes | | | 2 384 notices |
| | total PR | | | 2 835 notices |
| AN | Anonymes avant 1960 | | | |
| | cotes | 2 sur 1 000 | 474 311 cotes | 949 cotes |
| | titres (245/745/920 \$t) | 8 sur 1 000 | 473 109 zones titres | 3 785 titres |
| | barre de classement dans les titres | 8 sur 1 000 | 473 109 zones titres | 3 785 titres |
| | code langue | 20 sur 1 000 | 448 007 notices (MON) | 8 960 notices |
| | code pays | 20 sur 1 000 | 448 007 notices | 8 960 notices |

| | | | | |
|-----------|---|--|------------------|--------|
| | total AN hors cotes | | | 25 490 |
| | | | | 26 439 |
| | NOTICES DU LOT DE RATRAPAGE | = <u>anomalies non corrigées lors de la saisie</u> | | 50 000 |
| | total AN | | | 76 439 |
| CO | Publications de collectivités avant 1960 | | | |
| | cotes | 8 sur 1 000 | *160 681 notices | 1 285 |
| | vedettes collectivités/301/411) | 9 sur 1 000 | | 1 446 |
| | Titres (245) | 20 sur 1 000 | | 3 214 |
| | code langue | 20 sur 1 000 | | 3 214 |
| | code pays | 20 sur 1 000 | | 3 214 |
| | erreurs autres diverses | 90 sur 1 000 | | 14 461 |
| | total CO hors cotes | | | 25 549 |
| | total CO | | | 26 834 |
| AR | Actes royaux | | | |
| | cotes | 7 sur 1 000 | *42 043 notices | 294 |
| | vedettes auteurs et titre de forme (00/143/743) | 7 sur 1 000 | | 294 |
| | titres (245/748) | 12 sur 1 000 | | 505 |
| | codes langues | 15 sur 1 000 | | 631 |
| | codes pays | 15 sur 1 000 | | 505 |
| | erreurs autres diverses | 90 sur 1 000 | | 3 784 |
| | total AR hors cotes | | | 5 718 |
| | total AR | | | 6 012 |
| FB | Fonds du château de Fontainebleau | | | |
| | cotes | 8 sur 1 000 | *25 663 notices | 205 |
| | vedettes auteurs (100/400/700/920) | 8 sur 1 000 | | 205 |
| | titres (245/745/920 \$t) | 16 sur 1 000 | | 411 |
| | codes langues | 20 sur 1 000 | | 513 |
| | codes pays | 20 sur 1 000 | | 513 |
| | Erreurs autres diverses | 70 sur 1 000 | | 1 796 |
| | total FB hors cotes | | | 3 439 |
| | total FB | | | 3 644 |
| AA | Actes administratifs et judiciaires de l'Ancien Régime | | | |
| | cotes | 8 sur 1 000 | *110 857 notices | 887 |
| | auteurs personnes physiques et collectivités et titres factices (100/700/920 \$t/110/710/921/143/743) | 9 sur 1 000 | | 998 |
| | codes langue | 15 sur 1 000 | | 1 663 |
| | codes pays | 15 sur 1 000 | | 1 663 |
| | Erreurs diverses autres | 90 sur 1 000 | | 9 977 |
| | total AA hors cotes | 90 sur 1 000 | | 13 414 |
| | total AA | | | 14 301 |
| RE | Actes religieux de l'Ancien Régime | | | |
| | cotes | 8 sur 1 000 | *45 144 notices | 361 |
| | auteurs personnes physiques et collectivités, titres factices (100/700/920/110/710/921/143/743) | 9 sur 1 00 | | 406 |
| | codes langue | 15 sur 1 000 | | 677 |
| | codes pays | 15 sur 1 000 | | 903 |
| | erreurs autres diverses | 90 sur 1 000 | | 4 063 |
| | total RE hors cotes | | | 6 049 |
| | total RE | | | 6 410 |
| HL | Catalogue imprimé de l'Histoire de France | | | |

| | | | | |
|---------------------|---|-------------------------|------------------------------------|---------------|
| | cotes | 8 sur 1 000 | *227 624 notices | 1 821 |
| | auteurs personnes physiques et collectivités (100/700/110/710) | 9 sur 1 000 | | 2 049 |
| | titres (245/748/749) | 12 sur 1 000 | | 2 731 |
| | codes langue | 10 sur 1 000 | | 2 276 |
| | codes pays | 10 sur 1 000 | | 2 276 |
| | erreurs autres diverses | 90 sur 1 000 | | 20 486 |
| | total HL hors cotes | | | 29 818 |
| | total HL | | | 31 639 |
| HA | Histoire de France 1896-1936 | | | |
| | cotes | 4 sur 1 000 | *176 118 notices | 704 |
| | auteurs personnes physiques et collectivités (100/700/110/710) | 9 sur 1 000 | | 1 585 |
| | titre (245/748/749) | 20 sur 1 000 | | 3 522 |
| | codes langue | 10 sur 1 000 | | 1 761 |
| | codes pays | 10 sur 1 000 | | 1 761 |
| | erreurs autres diverses | 90 sur 1 000 | | 15 851 |
| | total HA hors cotes | | | 24 480 |
| | total HA | | | 25 184 |
| HB = HF = HP | Histoire de France 1936-1949 | | Histoire des familles Paris | |
| | cotes | 8 sur 1 000 | 42 509 cotes | 340 cotes |
| | auteurs personnes physiques et collectivités (100/700/110/710) | 9 sur 1 000 | 26 563 zones auteurs | 239 zones |
| | titres (245/748/749) | 16 sur 1 000 | 43 855 zones titres | 702 titres |
| | codes langue | 10 sur 1 000 | 39 642 notices | 396 notices |
| | codes pays | 10 sur 1 000 | 39 642 notices | 396 notices |
| | erreurs autres diverses | 90 sur 1 000 | 39 642 notices | 3 568 notices |
| | total HB+HF+HP hors cotes | | | 5 301 |
| | total HB+HF+HP | | | 5 641 |
| HC | Histoire de France 1950-1987 | | | |
| | cotes | 7 sur 1 000 | *27 731 notices | 194 |
| | auteurs personnes physiques et collectivités, titres factices (100/700/110/710)/143 | 7 sur 1 000 | | 194 |
| | titres (245/749) | 12 sur 1 000 | | 333 |
| | codes langue | 10 sur 1 000 | | 277 |
| | codes pays | 10 sur 1 000 | | 277 |
| | erreurs autres diverses | 90 sur 1 000 | | 2 496 |
| | total HC hors cotes | | | 3 577 |
| | total HC | | | 3 771 |
| AP | Actes parlementaires | | | |
| | cotes | 8 sur 1 000 | *39 598 notices | 317 |
| | auteurs personnes physiques et collectivités (100/700/710) | 9 sur 1 000 | | 356 |
| | titres (245) | 15 sur 1 000 | | 594 |
| | codes langue | 10 sur 1 000 | | 396 |
| | codes pays | 10 sur 1 000 | | 396 |
| | erreurs diverses autres | 90 sur 1 000 | | 3 564 |
| | total AP hors cotes | | | 5 306 |
| | total AP | | | 5 623 |
| VA | Catalogues de ventes d'objets d'art | | | |
| | cotes | 1 sur 1 000 (en fait 0) | *23 518 notices | 24 |

| | | | | |
|----------------|--|-------------------------|----------------------------|--------|
| | titres de forme (143/743) | 1 sur 1 000 | | 988 |
| | titres (245) | 42 sur 1 000 | | 988 |
| | codes langue | 42 sur 1 000 | | 988 |
| | codes pays | 2 sur 1 000 | | 47 |
| | autres erreurs diverses | 280 sur 1 000 | | 6 585 |
| | total VA hors cotes | | | 8 631 |
| | total VA | | | 8 655 |
| VB | Catalogues de ventes de livres | | | |
| | cotes | 1 sur 1 000 (en fait 0) | **53 849 notices | 54 |
| | titres de forme (143/743) | 3 sur 1 000 | | 162 |
| | titres (245) | 200 sur 1 000 | | 10 770 |
| | codes langue | 21 sur 1 000 | | 1 131 |
| | codes pays | 1 sur 1 000 | | 54 |
| | erreurs diverses autres | 200 sur 1 000 | | 10 770 |
| | total VB hors cotes | | | 22 886 |
| | total VB | | | 22 540 |
| BA + SL | Reliquat du Fonds Barrès | | Fonds Smith-Lesouëf | |
| | cotes | 4 sur 1 000 | **32 681 notices | 131 |
| | auteurs personnes physiques et collectivités, titres forgés (100/700/110/710/141/143/920) | 4 sur 1 000 | | 131 |
| | titres (245/745/749/920\$) | 11 sur 1 000 | | 359 |
| | codes langue | 14 sur 1 000 | | 458 |
| | codes pays | 14 sur 1 000 | | 458 |
| | erreurs autres diverses | 54 sur 1 000 | | 1 765 |
| | total BA+SL hors cotes | | | 3 170 |
| | total BA+SL | | | 3 301 |
| FN | Factums postérieurs à 1791 (recatalogués) | | | |
| | cotes | 1 sur 1 000 (en fait 0) | **34196 notices | 34 |
| | titres de forme et auteurs personnes physiques(143/743/100/700) | 3 sur 1 000 | | 103 |
| | titres (245) | 8 sur 1 000 | | 274 |
| | codes langue | 6 sur 1 000 | | 205 |
| | codes pays | 6 sur 1 000 | | 205 |
| | erreurs autres diverses | 100 sur 1 000 | | 3 420 |
| | total FN hors cotes | | | 4 206 |
| | total FN | | | 4 240 |
| FH | Factums postérieurs à 1791 (cartes à jouer) | | | |
| | cotes | 5 sur 1 000 | **32 000 notices | 160 |
| | titres de forme et auteurs personnes physiques (143/743/100/700) | 7 sur 1 000 | | 224 |
| | titres (245) | 170 sur 1 000 | | 5 400 |
| | codes langue | 26 sur 1 000 | | 832 |
| | codes pays | 26 sur 1 000 | | 832 |
| | erreurs diverses autres | 210 sur 1 000 | | 6 720 |
| | total FH hors cotes | | | 14 048 |
| | total FH | | | 14 208 |
| FA | Corde = Factums antérieurs à 1791 | | | |
| | cotes | 2 sur 1 000 | **42 000 notices | 84 |
| | titres de forme et auteurs personnes physiques (143/743/700) | 1 sur 1 000 | | 42 |
| | titres (245) | 5 sur 1 000 | | 210 |
| | codes langue | 3 sur 1 000 | | 126 |

| | | | | |
|----------------|---|--------------|------------------|----------------|
| | codes pays | 3 sur 1 000 | | 126 |
| | erreurs diverses autres | 80 sur 1 000 | | 3 360 |
| | total FA hors cotes | | | 3 864 |
| | total FA | | | 3 948 |
| TD + TL | Thèses de licence en droit françaises du 19e siècle | | | |
| | cotes | 1 sur 1 000 | **12 265 notices | 12 |
| | auteurs | 1 sur 1 000 | | 12 |
| | titres | 40 sur 1 000 | | 491 |
| | codes langue | 6 sur 1 000 | | 74 |
| | codes pays | 2 sur 1 000 | | 25 |
| | erreurs diverses autres | 30 sur 1 000 | | 368 |
| | total TD+TL hors cotes | | | 969 |
| | total TD+TL | | | 981 |
| TM | Thèses de médecine | | | |
| | cotes | 6 sur 1 000 | *102 012 notices | 612 |
| | auteurs | 4 sur 1 000 | | 408 |
| | titres | 9 sur 1 000 | | 918 |
| | codes langue | 6 sur 1 000 | | 612 |
| | codes pays | 6 sur 1 000 | | 612 |
| | erreurs diverses autres | 45 sur 1 000 | | 4 590 |
| | total TM hors cotes | | | 7 140 |
| | total TM | | | 7 752 |
| LI | Liturgie | | | |
| | cotes | 6 sur 1 000 | **20 000 notices | 120 |
| | titres factices, auteurs personnes physiques et collectivités (141/741/143/743/7 00/710) | 5 sur 1 000 | | 100 |
| | titres | 11 sur 1 000 | | 220 |
| | codes langue | 14 sur 1 000 | | 280 |
| | codes pays | 14 sur 1 000 | | 280 |
| | Erreurs autres diverses | 6 sur 1 000 | | 1 200 |
| | total LI hors cotes | | | 2 080 |
| | total LI | | | 2 200 |
| | TOTAL COTES | | | 15 608 |
| | TOTAL VEDETTES | | | 33 978 |
| | TOTAL HORS COTES ET VEDETTES | | | 468 301 |
| | TOTAL GENERAL | | | 517 887 |

3. Cotes en erreur

COTES EN ERREUR

On a utilisé par ordre de priorité décroissant :
 - les statistiques de la V1
 - le nombres de notices saisies après dédoublemnages pré-migration V1: *
 - le nombre de notices à saisir, estimé : **

| Catalogue | Nombre Total de cotes ou de notices | % de l'épine dorsale constituée | % à reconstituer | Nombre de cotes à vérifier |
|-------------|-------------------------------------|---------------------------------|------------------|----------------------------|
| CG | 1 936 097 | 90,2 | 9,8 | 189 738 |
| SU | 1 018 205 | 90,5 | 9,5 | 96 729 |
| LX | 330 155 | 84,5 | 15,5 | 51 174 |
| PR | 225 420 | 88,3 | 11,7 | 26 374 |
| AN | 474 311 | 83,5 | 16,5 | 78 261 |
| CO | *160 681 | 77,8 | 22,2 | |
| AR | *42 043 | 91,8 | 8,2 | |
| FB | *25 663 | 26,5 | 73,5 | |
| AA | *110 857 | 88,6 | 11,4 | |
| RE | *45 144 | 88,6 | 11,4 | |
| HA | *176 118 | 94,3 | 5,7 | |
| HB +HF + HP | 42 509 | 94,3 | 5,7 | 2 423 |

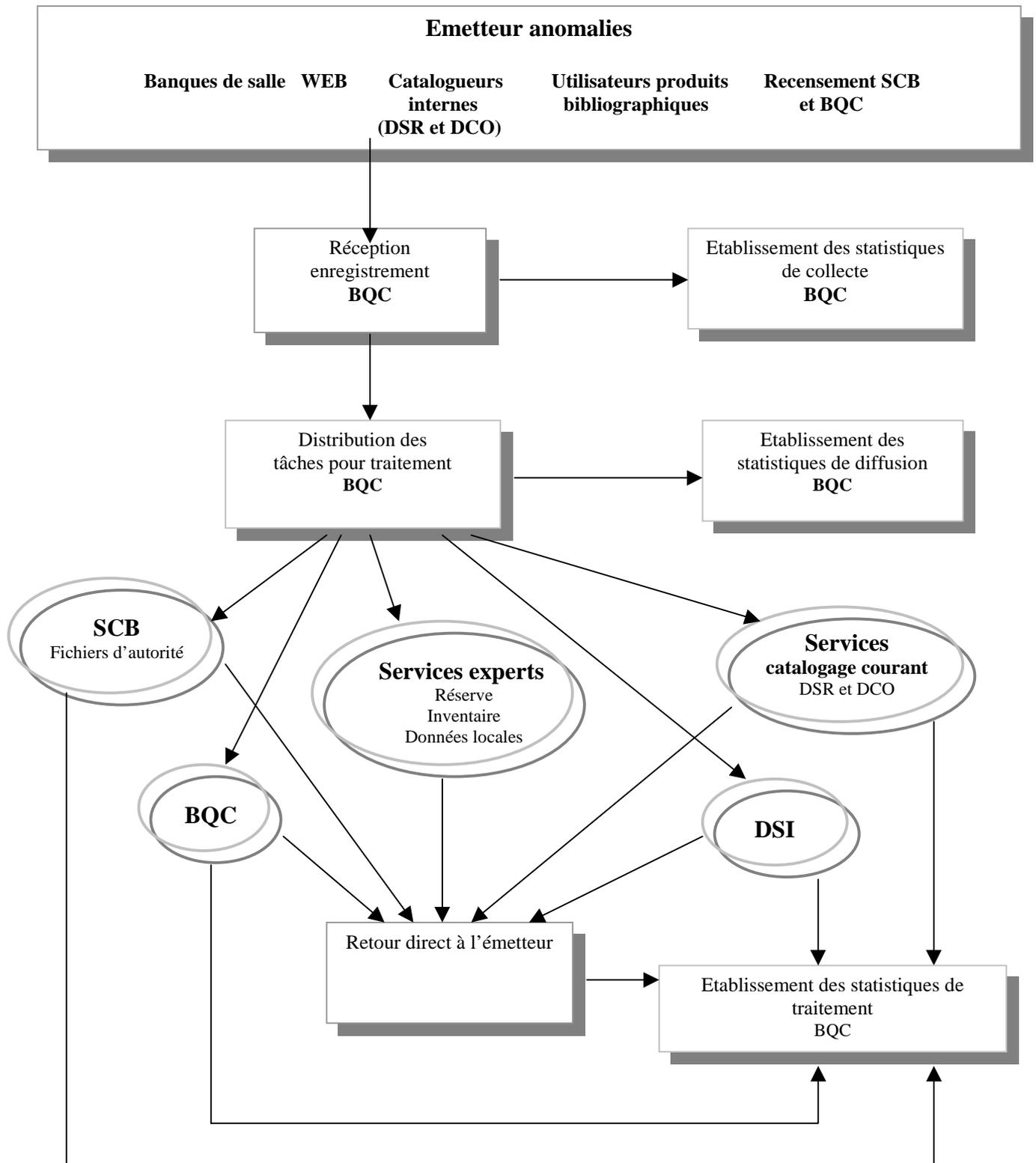
Fichiers à charger en V2 ou plus tard : on ignore quel sera le taux de constitution de l'épine dorsale

| | | | | |
|---------------------------------------|----------|--|--|----------------|
| HL | *227 624 | | | |
| HC | *27 731 | | | |
| AP | *39 598 | | | |
| VA | *23 518 | | | |
| VB | **53 849 | | | |
| BA | **13 430 | | | |
| SL | **19 251 | | | |
| FN | **34196 | | | |
| FH | **32 000 | | | |
| FA | **42 000 | | | |
| TD+TL | **12 265 | | | |
| TM | *101 933 | | | |
| LI | **20 000 | | | |
| Total à corriger dès 2001 (V2) | | | | 444 699 |

4. Bureau qualité du catalogue (BQC) : panorama des fonctions

BUREAU QUALITE DU CATALOGUE (BQC)

Panorama des fonctions



5. Exemples d'index du cédérom *Catalogue général des imprimés des origines à 1970*
 - Ecran : « Feuilletage auteur » : recherche à Balzac où l'on voit les différentes formes à uniformiser et l'incidence d'une erreur de frappe :

Chisissez une entrée dans cette liste

Feuilletage

Catalogue général de la BnF

- Ecran du feuilletage par titre à LE bulletin où l'on voit un exemple d'absence de

Chisissez une entrée dans cette liste

Feuilletage

Catalogue général de la BnF

barres de classement après l'article :
 6. Exemple de grappe issue du SI : Balzac
 - Notices ayant pour vedette : Balzac

BN-OPALE PLUS Index 15/09/1999 09:06

Nouvelle recherche Parcours recherche Liste des demandes

Au moins 50 auteurs répondent à votre recherche
 Sélectionnez un ou plusieurs auteurs

Rappel des critères de recherche

| Auteur | Notices |
|--|---------|
| Balz-Cochois, Helgard | 1 |
| Balza, José (1939-....) | 10 |
| ▼ Balzac | 72 |
| Imprimés / Cat. Général Auteurs (jusqu'en 1959) | 2 |
| Imprimés / Cat. Général Auteurs - Supplément (jusqu'en 1959) | 50 |
| Imprimés / Cat. Général des ouvrages anonymes (jusqu'en 1959) | 6 |
| Imprimés / Cat. Général des collectivités Auteurs (jusqu'en 1959) | 1 |
| Imprimés / Cat. Fonds Fontainebleau (jusqu'en 1936) | 4 |
| Imprimés / Cat. de l'Histoire de France (1895-1987) | 1 |
| Imprimés / Cat. de l'Histoire de France (1895-1987) | 2 |
| Fonds Smith-Lesouëf | 6 |
| Balzac | 1 |
| Balzac, Anne Charlotte Laure Sallambier (Mme Bernard François de) | 1 |
| Balzac, Anne Charlotte Laure Sallanibier (Mme Bernard François de) | 1 |
| ▶ Balzac, B de | 2 |

Annuler la sélection

Imprimer

Complétez ou modifiez : balzac ... et relancez

Voir les notices

Cliquez sur une entrée pour la sélectionner ou sur une icône pour avoir plus d'information sur l'entrée correspondante

- Notices ayant pour vedette : Balzac, H., Honoré, Honoré de, etc. plus loin dans le même index.

BN-OPALE PLUS Index 15/09/1999 09:24

Nouvelle recherche Recherches précédentes Parcours recherche Liste des demandes

Au moins 50 auteurs répondent à votre recherche
 Sélectionnez un ou plusieurs auteurs

Rappel des critères de recherche

| Auteur | Notices |
|---|---------|
| Balzac, H | 1 |
| ▼ Balzac, H de | 21 |
| Imprimés / Cat. Général Auteurs - Supplément (jusqu'en 1959) | 6 |
| Imprimés / Cat. Général des ouvrages anonymes (jusqu'en 1959) | 1 |
| Imprimés / Cat. Fonds Fontainebleau (jusqu'en 1936) | 2 |
| Imprimés / Cat. de l'Histoire de France (1895-1987) | 1 |
| Imprimés / Cat. de l'Histoire de France (1895-1987) | 1 |
| Fonds Smith-Lesouëf | 10 |
| Balzac, Honoré | 1 |
| ▼ Balzac, Honoré de | ? |
| Imprimés / Cat. Général Auteurs et anonymes (1960-1970) | 381 |
| Imprimés / Cat. Général Auteurs (jusqu'en 1959) | 518 |
| Imprimés / Cat. Général Auteurs - Supplément (jusqu'en 1959) | ? |
| Imprimés / Cat. Général Auteurs - Supplément (jusqu'en 1959) | 1 |
| Imprimés / Cat. Général des ouvrages anonymes (jusqu'en 1959) | 1 |

Annuler la sélection

Imprimer

Complétez ou modifiez : balzac ... et relancez

Voir les notices

Cliquez sur une entrée pour la sélectionner ou sur une icône pour avoir plus d'information sur l'entrée correspondante

BIBLIOGRAPHIE

Les références bibliographiques sont présentées selon la dernière norme de transcription Z 44-005⁴⁶.

Les documents consultés à des titres divers, ont été regroupés et classés selon trois subdivisions par ordre de priorité afin de faciliter l'accès à cette bibliographie, de la rendre la plus lisible possible :

- d'abord le thème principal de la référence auquel on peut la rattacher.
- puis l'ordre alphabétique des auteurs

Trois principaux thèmes ont été retenus pour le classement des outils. Leur ordre de succession se veut pertinent et logique : il va du général au plus particulier :

- Bibliothèque nationale de France, Direction des services et des réseaux, Agence bibliographique nationale, expériences extérieures (stage à la BDIC)
 - Catalogue BN-Opale et BN-Opale Plus, migration des données
 - Contrôle de la qualité, nettoyage et dédoublonnage, recherche d'erreurs, cohérence de systèmes d'information
- Bibliothèque nationale de France, Direction des services et des réseaux, Agence bibliographique nationale, expériences extérieures (stage à la BDIC)

- « Décret n° 94-3 du 3 janvier 1994 : portant création de la Bibliothèque nationale de France », *Journal officiel de la République française*, 4 janvier 1994, p. 149
- « La BDIC à l'aube du XXI^e siècle », *Matériaux pour l'histoire de notre temps*, janvier - juin 1998, n° spécial 49-50
- « Le système d'information de la Bibliothèque nationale de France », *Trajectoire*, juin 1998, n° spécial, 19 p. : ill. en noir et en coul., couv. ill. en coul. ; 30 cm.
- « Les fonds nouveaux à la BDIC et au Musée d'histoire contemporaine », *Matériaux pour l'histoire de notre temps*, juillet - septembre 1991, n° 24
- « Visage de la BnF : [organigramme] », février 1999, *Trajectoire*, n° 25, p. 4
- BIBLIOTHEQUE DE DOCUMENTATION ET D'INFORMATION CONTEMPORAINE. *Bibliothèque de Documentation Internationale Contemporaine. Bibliothèque de Documentation Internationale Contemporaine - Spécialisée en histoire contemporaine, relations internationales et guerres mondiales / BDIC*. Accessible via Internet, URL : <http://www.paris10-uparis/bdic>
- BIBLIOTHEQUE DE DOCUMENTATION ET D'INFORMATION CONTEMPORAINE. *La veille documentaire à la BDIC : réflexion sur les activités de la BDIC : objectifs et mise en oeuvre*. - Nanterre : BDIC, 1997. - Non paginé : ill. ; 30 cm. - (Collection des publications de la BDIC). ISBN 2-901658-57-1 (br.) : 35 F
- BIBLIOTHEQUE NATIONALE DE FRANCE. *L'informatique à la Bibliothèque nationale de France en 1995*. - Paris : Bibliothèque nationale de France. - Dépliant
- BIBLIOTHEQUE NATIONALE DE FRANCE. *Réorganisation générale des services [: document interne]*. - [Paris : Bibliothèque nationale], 1998. - 35 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. Service de coordination bibliographique. *Le catalogue informatisé au 1^{er} avril 1995*. - Paris : Bibliothèque nationale de France. - Dépliant
- BIBLIOTHEQUE NATIONALE. *Centre de coordination bibliographique et technique*. - Paris : Bibliothèque nationale. - Dépliant
- BLASSELLE, Bruno. « Classement des savoirs et classement des livres de la Bibliothèque nationale de France ». In *Tous les savoirs du monde : encyclopédies et bibliothèques, de Sumer au XXI^e siècle / sous la dir. de Roland Schaer*. - Paris : Bibliothèque nationale de France : Flammarion, 1996. - 495 p. : ill. en noir et en coul., couv. ill. en coul. ; 30 cm. ISBN 2-7177-1977-6 (Bibliothèque nationale de France). - ISBN 2-08-010217-6 (Flammarion) (rel.) : 450 F. P. 194-199
- BOUDET, Isabelle. « Agence bibliographique nationale française : tradition et innovation », *International cataloguing and bibliographic control*, avril - juin 1996, vol. 25, n° 2, p.34-37
- FINELLI-LEMELLE, Françoise. « Le centre de coordination bibliographique et technique », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 4^e trimestre 1986, n° 133, p. 12-13
- HUE, Joseph. *1917-2000 : BDIC, Bibliothèque de documentation internationale contemporaine : un organisme public d'information et de recherche international / présentée par Joseph Hüe,...* - Nanterre : Paris : BDIC, 1997. - 128 p. : ill. en

⁴⁶ *Documentation : références bibliographiques : contenu, forme et structure : Z 44-005 : décembre 1987 / Association française de normalisation*. - Paris : AFNOR, 1988. - 13 p. ; 24 cm. - (ISSN 0335-3931).

noir et en coul. ; 30 cm. - (Collection des publications de la BDIC). ISBN 2-901658-56-2 (br.) : 150 F

- KAPLAN, Hélène. « Documents sans frontières : la B.D.I.C. », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 3^e trimestre 1986, n° 132, p. 23-25
- PION, Romulad. « Lecteurs, qui êtes-vous ? », *Chroniques de la bibliothèque nationale de France*, octobre - novembre - décembre 1999, n° 8, p. 3-4
- PORTES, Laurent. « L'ordre des livres selon Nicolas Clément ». In *Mélanges autour de l'histoire des livres imprimés et périodiques*. - Paris : Ed. de la Bibliothèque nationale de France, 1999, p. 64-92
- SANTANTONIOS, Laurence. « Embellie de l'été à la BnF », *Livres-Hebdo*, n° 349, 17 septembre 1999, p. 56
- VOUILLOT, Bernard. « 160 ans d'inventaire des livres imprimés de la Bibliothèque nationale : de la carte à jouer au CD-Rom : 1837-1997 ». In *Mélanges autour de l'histoire des livres imprimés et périodiques*. - Paris : Ed. de la Bibliothèque nationale de France, 1999. - P. 267-293

- Catalogue BN-Opale et BN-Opale Plus, migration des données

- ALESTCHENKOFF, Boris. « BN-OPALE PLUS : le catalogue en ligne », *Chroniques de la Bibliothèque nationale de France*, mai - juin 1999, n° 6, p. 4
- BEAUDIQUEZ, Marcelle, BEAUGENDRE, Anne-Marie. « Du tiroir à l'écran : les opérations de conversion rétrospective de la Bibliothèque nationale de France », *International cataloguing and bibliographic control*, october - december 1997, vol. 26, n° 4, p. 85-90
- BEAUDIQUEZ, Marcelle. « La conversion rétrospective en chiffres », *Nouvelle(s) BN : information des personnels de la Bibliothèque nationale*, septembre - octobre 1992, n° 1, p. 7
- BEAUDIQUEZ, Marcelle. « La conversion rétrospective : geste historique au service de la mémoire et mission de service public, la conversion rétrospective est aussi l'œuvre d'une équipe », *Trajectoire*, mars 1993, n° 3, [1] p.
- BEAUDIQUEZ, Marcelle. « Le chantier conversion rétrospective de la Bibliothèque nationale », *Bulletin des bibliothèques de France*, 1993, t. 38, n° 3, p. 8-19
- BEAUGENDRE, Anne-Marie, EHLING, Thérèse. « La conversion rétrospective à la portée de tous », *La lettre [de la Bibliothèque nationale]*, avril 1991, n° 7, p. 14-17
- BEAUGENDRE, Anne-Marie. « Le Catalogue général des imprimés des origines à 1970 », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 1^{er} trimestre 1997, n° 174, p. 15-16
- BERNARD, Annick. « Conversion rétrospective - Conversion « autonome » : d'après l'expérience de la Bibliothèque nationale », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 1^{er} trimestre 1988, n° 138, p. 11-12
- BIBLIOTHEQUE NATIONALE DE FRANCE. *BN-OPALE PLUS : Le catalogue des collections patrimoniales (imprimés)*. - Accessible via Internet : URL : <<http://catalogue.bnf.fr>>
- BIBLIOTHEQUE NATIONALE DE FRANCE. Direction du Développement Scientifique et des Réseaux. *Conversion rétrospective des catalogues à la BnF : CRI [: document interne] / Direction du Développement Scientifique et des Réseaux, Service des Chantiers de Conversion Rétrospective*. - [Paris] : [Bibliothèque nationale de France], 1998. - 4 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. Direction du Développement Scientifique et des Réseaux. *Conversion rétrospective : noms codés des catalogues et fichiers [: document interne] / Direction du Développement Scientifique et des Réseaux, Service des Chantiers de Conversion Rétrospective*. - [Paris] : [Bibliothèque nationale de France], 1998. - 3 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. Direction du Développement Scientifique et des Réseaux. *Etat au 31 décembre 1996 [de la Conversion Rétrospective : document interne] / Direction du Développement Scientifique et des Réseaux, Service des Chantiers de Conversion Rétrospective*. - [Paris] : [Bibliothèque nationale de France], [1996]. - 12 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. *La migration des données issues de Sycomore [: document interne] / Chantal Weill ; avec la collab. d'Annie Derrez pour les aspects techniques*. - [Paris] : [Bibliothèque nationale de France], 1999. - 5 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. Service des chantiers de la conversion rétrospective. « Conversion rétrospective des catalogues de la Bibliothèque nationale de France (CRI) et chargements dans BN-Opale + [: document interne]. - [Paris] : [Bibliothèque nationale de France], 1999. - 4 p. ; 30 cm.
- BIBLIOTHEQUE NATIONALE DE FRANCE. *Synthèse conversion V1 : note de projet [: document interne] / D. Lévêque*. - [Paris] : Cap Gemini, 1997. - 19 p. ; 30 cm.

- BOUDET, Isabelle, CLOAREC, Thierry. « Projet d'enrichissement pour le catalogue de la Bibliothèque de France », *Bulletin des bibliothèques de France*, 1993, t. 38, n° 3, p. 50-52
- CONFEDERATION GENERALE DU TRAVAIL. Section de la Bibliothèque nationale de France. *A propos du Système informatique de la Bibliothèque nationale de France : dérapage d'un projet et pilotage en dérive* / CFDT Culture, Section de la Bibliothèque nationale de France. - [Paris] : [Bibliothèque nationale de France], 1999. - 22 p. ; 30 cm.
- EHLING-BEGUE, Thérèse. *Conversion d'un ensemble de fichiers et d'un catalogue (environ 400 000 notices) dont les problèmes sont la création de points d'accès multiples et une cotation souvent incomplète et même absente* [: *Projet Professionnel Personnel*]. - Villeurbanne : Institut de Formation des Bibliothécaires, 1992. - 10 p. ; 30 cm.
- NAUDÉ, Gabriel. *Advis pour dresser une bibliothèque. Précédé de L'advis, manifeste de la bibliothèque érudite* / par Claude Jolly, ... - [Reprod. en fac-sim.]. - Paris : Aux Amateurs de livres, 1990. - XXIV-164 p. ; 18 cm. Fac-sim. de l'éd. de Paris : Rolet Le Duc, 1644. - Bibliogr. p. 23-24. - ISBN 2-252-02730-4 (br)
- SOLOMIAC, Jacqueline. « La conversion rétrospective à la Bibliothèque nationale », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 4^e trimestre 1991, n° 153, p. 20-22

- Contrôle de la qualité des systèmes d'information et des catalogues, nettoyage et dédoublonnage, recherche d'erreurs, cohérence de systèmes d'information

- ASSOCIATION FRANÇAISE DE NORMALISATION. *Management de la qualité et assurance de la qualité - Vocabulaire : NF EN ISO 8402 Juillet 1995 / AFNOR*. - Paris : AFNOR, 1995. - 43 p. ; 30 cm. Remplace NF X50-120-1987 (X 50-120). - (Br.) : 405,00 F
- AYRES, F. H. « Duplicates and other manifestations : a new approach to the presentation of bibliographic information », *Journal of librarianship*, october 1990, vol. 22, n° 4, p. 236-251
- BITTON, Dina. « Duplicate record elimination in large data files », *ACM transactions on database systems*, 1983, vol. 8, n° 2, p. 255-265
- CHAMOIS-CHAYNES, Jacqueline. *Gestion d'une base bibliographique informatisée au S.C.D. Lyon 2 : contrôle de qualité et de cohérence : Projet Professionnel Personnel...* - Villeurbanne : Institut de formation des bibliothécaires, 1993. - 10 p. ; 30 cm.
- *Contrôler la qualité et la cohérence d'un catalogue / sous la dir. de Marie - Elise Fréon*. - Villeurbanne : Institut de formation des bibliothécaires, 1996 (69-Villeurbanne : Impr. Terreaux). - 171 p. : couv. ill. ; 21 cm. - (Collection La boîte à outils, ISSN 1259-4857 ; 4). Notes bibliogr. Glossaire. - ISBN 2-910966-03-8 (br.) : 150 F
- DESRICHARD, Yves. « Le dédoublonnage des banques de données bibliographiques : un état de l'art », *Documentaliste - Sciences de l'information*, vol. 34, n° 2, 1997, p. 82-89 ; accessible via Internet, URL : <http://www.adbs.fr/adbs/prodserv/docu/003/html/1AI00212.htm>
- GOYAL, Pankaj. « An investigation of different string coding methods », *Journal of the American society for information science*, 1984, vol. 35, n° 4, p. 248-252
- GOYAL, Pankaj. « Duplicate record identification in bibliographic databases », *Information systems (Oxford)*, 1987, vol. 12, n° 3, p. 239-242
- GUY, Michèle. *Rapport annuel APP : 1998* [: *document interne*] / Michèle Guy. - [Paris] : [Bibliothèque nationale de France], 1998. - 13 p.
- HICKEY, Thomas B., RYPKA, David J. « Automatic detection of duplicate monographic records », *Journal of library automation*, june 1979, vol. 12, n° 2, p. 125-142
- KILGOUR, F. G. « Retrieval of single entries from a computerised library catalog file », *Proceedings of the American society for information science*, 1968, vol. 5, p. 133-136
- LAHARY, Dominique. « La gestion des accès aux notices récupérées : ou comment digérer convenablement les importations », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 2^e trimestre 1994, n° 163, p. 137-139
- McLAURY, Keith D. « Automatic merging of monographic databases - uses of fixed-length keys derived from title strings », *Journal of library automation*, 1979, vol. 12, n° 2, p. 143-155
- MONTEILLET, Françoise. *Le contrôle de qualité du catalogue signalétique à la Bibliographie nationale française - Livres : étude des moyens existants à mettre en œuvre pour optimiser la recherche bibliographique* [: *Projet Professionnel Personnel*]. - Villeurbanne : Institut national de formation des bibliothécaires, 1993. - 14 p. ; 30 cm.
- O'NEILL, Edward T. « Characteristics of duplicate records in OCLC's online union catalog », *Library resources and technical services*, 1993, vol. 37, n° 1, p. 59-71

- O'NEILL, Edward T., VIZINE GOETZ, D. « Quality control in online databases », *Annual review of information science and technology*, 1988, vol. 23, p. 125-156
- *Offrir aux publics un catalogue en ligne* / Institut de Formation des Bibliothécaires ; sous la dir. d'Eliane Bernhart. - Villeurbanne : IFB, 1996. - (Collection La boîte à outils, 2).
- ONORATO, Eveline S., BIANCHI, C. « Automatic identification of duplicates after multidatabase online searching », *Online review*, december 1981, vol. 5, n° 6, p. 445-451
- RENOULT, Daniel. « L'informatique au service du public : de la Bibliothèque nationale à la Bibliothèque nationale de France », *Bulletin d'informations de l'Association des Bibliothécaires Français*, 1^{er} trimestre 1997, n° 174, p. 85-90
- RIDLEY, M. J. « An expert system for quality control and duplicate detection in bibliographic databases », *Program*, 1992, vol. 26, n° 1, p. 1-18
- ROGER, Danielle. « Catalogues en ligne accessibles par le public : recherche exploratoire », *Bulletin des bibliothèques de France*, 1994, t. 39, n° 2, p. 43-50
- SLACH, June E. « Detection and elimination of duplicates from multidatabases searches », *Bulletin of the medical library association*, july 1985, vol. 73, n° 3, p. 235-237
- TONEY, Stephen R. « Cleaning and deduplication of an international bibliographic database », *BCIN*, 1992, p. 19-28 ; repris dans *Information technology and libraries*, 1992, vol. 11, n° 1, p. 19-28 1994, vol. 11, n° 1, p. 25
- VAYSSADE, Claire, GUITTON, Janine. « La démarche qualité à l'Agence bibliographique nationale », *Bulletin des bibliothèques de France*, 1998, t. 43, n° 1, p. 35-38
- WANNINGER, P. D. « Is the OCLC database too large ? A study of the effects of duplicate records in the OCLC system », *Library resources and technical services*, 1982, vol. 26, p. 353-361
- WILLIAMS, Martha E., MACLAURY, Keith D. « Automatic merging of monographic data bases - identification of duplicate records in multiple files : the IUCS scheme », *Journal of library automation*, june 1979, vol. 12, n° 2, p. 156-168
- YANNAKOUDAKIS, E. J. « Towards a universal record identification and retrieval scheme », *Journal of informatics*, 1979, vol. 3, p. 7-11
- YANNAKOUDAKIS, E. J., AYRES, F. H., HUGGILL, J. A. « Matching of citations between non-standardized databases », *Journal of the American society for information science*, 1990, vol. 41, n° 8, p. 599-610
- YANNAKOUDAKIS, E. J., AYRES, F. H., HUGGILL, J. A. W. « Character coding for bibliographic record control », *The computer journal*, 1979, vol. 13, n° 1, p. 53-60