

0899

1988  
Stages  
30

Université Lyon I  
DESS Informatique Documentaire

ETUDE DE DIFFERENTS SYSTEMES D'INDEXATION ELABORES  
A PARTIR DU REPERTOIRE LAVAL

VILLIERS, Béatrice

STAGE : Bibliothèque Nationale  
Service Informatique et Organisation

Villeurbanne, septembre 1988

1988  
Stages  
30

Université Lyon I  
DESS Informatique Documentaire

ETUDE DE DIFFERENTS SYSTEMES D'INDEXATION ELABORES  
A PARTIR DU REPERTOIRE LAVAL

VILLIERS, Béatrice

STAGE : Bibliothèque Nationale  
Service Informatique et Organisation

Villeurbanne, septembre 1988

## INTRODUCTION.

La Bibliothèque Nationale est un établissement public à caractère administratif, placé sous la tutelle du Ministère de la Culture et de la Communication.

L'essentiel de ses collections provient du dépôt légal. Celui-ci, instauré par l'ordonnance de Montpellier de 1537, par François 1er, visait à l'époque les livres imprimés. Puis, avec la multiplication des productions éditoriales, il s'est appliqué à d'autres types de documents : autres les périodiques, les estampes, affiches, photos, cartes et plans, partitions, cassettes, disques, vidéo, films.

De plus, la BN s'est enrichie de manuscrits et livres anciens, au cours des siècles. Elle possède enfin un fonds important de médailles et antiques.

Elle complète également ses collections par le biais d'achats, échanges ou dons de livres dans les domaines de la littérature et des sciences humaines.



1988  
Stages  
30

La BN possède donc des collections énormes et de contenu très divers, le dépôt légal couvrant tous les domaines de la connaissance.

Son rôle premier est la collecte des documents soumis au dépôt légal. Elle doit veiller à ce que les éditeurs respectent cette loi. Puis, elle assure le traitement de ces documents, travail d'un ampleur considérable. Enfin, la BN exploite ces documents et les communique au public, depuis 1692. Elle les conserve ensuite dans les meilleures conditions possibles.

Cette conservation pose des problèmes de plusieurs natures : le stockage nécessite des kilomètres de rayonnages. la BN n'a cessé de s'étendre, depuis le début du siècle, tout d'abord dans le quadrilatère de la rue Richelieu, puis l'annexe de Versailles, le château de Sablé et enfin depuis 1985 l'immeuble de la rue Vivienne a été terminé. Tous ces locaux ne sont pas suffisants face à l'accroissement perpétuel des collections. De plus, cette conservation se heurte aux problèmes de l'autodestruction du papier par acidification.

La BN, a de nombreux a donc entrepris des recherches en matière de conservation, et utilise les nouvelles technologies.

Le nouveau projet de la "Grande Bibliothèque" annoncé par le gouvernement devrait résoudre certains de ces problèmes.

Nous avons réalisé notre stage au Service Informatique et Organisation, sous la direction de Monsieur Serge Salomon et de madame Elisabeth Freyre. Le service comprend une vingtaine de personnes. Il a réalisé à la fois l'informatisation du traitement de documents et celle de différentes opérations de gestion.

Il assure la gestion du service comptable et financier de la BN, grâce à deux ordinateurs BULL. Il gère aussi la vente des éditions, par micro ordinateur.

Le traitement de documents imprimés se fait directement en ligne depuis le premier janvier 1988. La base bibliographique des imprimés de la BN, BN-OPALE, contient 570000 notices de livres et périodiques, avec un accroissement annuel de 70000 notices. Elle contient tous les livres français depuis 1975.

Les catalogues bibliographiques des départements spécialisés sont gérés dans des bases distinctes, grâce à un logiciel maison, OPALINE, implanté sur un BULL DPS7.

Actuellement les projets du SIO sont nombreux. Parmi ceux-ci, l'informatisation du traitement des documents continue à être étendue à d'autres départements; et la mise en place d'un réseau national d'indexation RAMEAU nécessite la gestion de bandes de vedettes matières.

Le stage a été divisé en deux périodes. Pendant la première partie, nous avons participé à l'élaboration d'un nouveau format adapté aux manuscrits enluminés. Par la suite, nous nous sommes penchées sur le problème de l'indexation matière.

ETUDE DE DIFFERENTS SYSTEMES D'INDEXATION MATIERE  
ELABORES A PARTIR DU REPERTOIRE LAVAL.

1. CHAMP DE L'ETUDE

L'étude que nous devons réaliser a pour objet de faire le point sur les différents systèmes d'indexation matière qui existent actuellement.

L'indexation matière représente une partie très importante du traitement des documents. Elle vise le contenu des documents et permet aux utilisateurs de les retrouver. Elle prend beaucoup de temps aux bibliothécaires. Elle leur pose des problèmes car elle peut être envisagée de façons très diverses, qui oscillent entre les deux extrêmes que sont l'indexation analytique et l'indexation systématique. Le choix d'une politique d'indexation est subjectif. Pourtant, dans un système documentaire, il est indispensable de maintenir une certaine cohérence, ce qui n'est pas facile, surtout lorsqu'il y a plusieurs indexeurs.

Nous avons commencé par faire une recherche bibliographique. Nous avons pu constater qu'il existe peu de documents concernant le sujet.

Nous avons donc décidé d'aller directement à la source, voir concrètement comment fonctionnent les systèmes de quelques bibliothèques et centres de documentation qui nous ont semblé être représentatifs des diversités du monde de la documentation, à savoir : les grandes bibliothèques d'envergure nationale telles que la Bibliothèque Nationale (BN), la Bibliothèque Publique d'Information (BPI); des bibliothèques plus petites comme la Bibliothèque Municipale (BM) de la Part-Dieu, des bibliothèques d'organismes dont le rôle documentaire est primordial : l'UNESCO, l'Institut du Monde Arabe (IMA), le Centre de documentation Sciences humaines (CDSH) du CNRS, la Médiathèque de la Villette et le Cercle de la Librairie. Nous sommes également allées dans une bibliothèque de grande école, celle de l'Ecole Polytechnique.

Il nous a semblé indispensable d'interroger les responsables de l'indexation matière de ces organismes. Pour cela nous avons donc élaboré un questionnaire afin d'avoir une plus grande cohérence dans les réponses, de pouvoir réaliser une comparaison entre les différents systèmes. (cf. Annexe)

Pour ce questionnaire, nous avons voulu d'abord faire le point sur le vocabulaire contrôlé utilisé, sur le plan historique et sur le plan syntaxique. Puis nous avons vu la gestion de ce vocabulaire, notamment sa mise à jour.

Le deuxième point du questionnaire concerne l'utilisation de ce système par ses gestionnaires , et par les "clients" de la bibliothèque ou du centre de documentation. Nous avons voulu connaître d'une part l'utilisation qui en est faite par les indexeurs, et d'autre part la façon dont le système est interrogé.

A travers ce questionnaire, nous avons essayé de percevoir le degré de satisfaction à la fois des indexeurs et du public.

Enfin, il nous a semblé intéressant de compléter nos investigations en allant interroger le producteur du logiciel Alexis, la Société ERLI.

## 2. PRESENTATION DES ETABLISSEMENTS

Les organismes documentaires visités révèlent une grande diversité. D'une part, les fonds sont divergents par leurs volumes et par les disciplines qu'ils couvrent. D'autre part, ces organismes ont des fonctions très variées, ne ciblent pas les mêmes publics, et ils se situent dans des contextes très différents.

Les bibliothèques à caractère national, la BN et la BPI, possèdent toutes les deux des fonds encyclopédiques, mais leurs points de vue divergent. La BN dont les fonds sont constitués par le dépôt légal, joue historiquement un rôle de bibliothèque de conservation, et subit les contraintes politiques, culturelles et matérielles qui en découlent. Mais à l'heure actuelle, elle doit aussi permettre au public d'accéder le plus facilement possible à l'information. De plus, à notre époque où l'information se généralise, la BN se voit confier un nouveau rôle; celui de moteur des autres bibliothèques en matière de coopération et de normalisation du traitement des documents.

La BPI, quant à elle, est récente et constitue avant tout une bibliothèque d'actualité. Son rôle essentiel est de mettre l'information à la disposition d'un large public. Elle sert de modèle aux plus petites bibliothèques, par son aspect "moderne", privilégiant le libre accès aux documents de toutes natures. Elle partage avec la BN le rôle de coordinateur de la normalisation au plan national.

Ces deux grandes bibliothèques servent de référence pour la majorité des autres établissements, d'une part les bibliothèques d'étude et d'autre part les bibliothèques dites de lecture public.

Parmi celles-ci, la BM de la Part-Dieu à Lyon, frappe par son dynamisme et son souci de coopération. La bibliothèque de l'Ecole Polytechnique fait partie des bibliothèques d'étude tout en s'adressant à un public bien particulier et restreint, constitué par les élèves de l'école et les chercheurs des laboratoires du site de Palaiseau.

La Médiathèque de la Villette se situe en marge des autres, par sa spécialisation dans le domaine des sciences, techniques et industrie. Rattachée à la Cité des Sciences et de l'Industrie, elle poursuit le même but d'ouverture au grand public. La Médiathèque "tout public" s'adresse, comme son nom l'indique, à tous, du premier âge jusqu'au niveau d'étude universitaire, et la Médiathèque "spécialisée" est susceptible d'intéresser un public de chercheurs.

Deux grandes bibliothèques, celle de l'UNESCO et celle de l'Institut du Monde Arabe, possèdent des fonds à première vue encyclopédiques. Mais elles ne peuvent pas être assimilées aux bibliothèques traditionnelles car elles dépassent le cadre national français et leur politique est celle de l'institution à laquelle elles sont rattachées; l'UNESCO s'intéresse tout particulièrement à l'éducation, les sciences et la culture, notamment dans les pays du tiers-monde, l'IMA vise à mieux faire connaître la culture arabe aux occidentaux.

De même, le CDSH a pour fonction de développer les services rendus par le CNRS en matière d'information dans le domaine des sciences humaines et sociales. Il produit une vingtaine de banques de données, dont chacune couvre un domaine précis. Les banques de données sont gérées par des spécialistes du domaine, indépendamment les unes des autres, en gardant toutes un niveau scientifique.

Le dernier système étudié, celui du Cercle de la Librairie, ne traite pas les documents dans la même optique que les autres organismes documentaires. Il gère l'ensemble de la production éditoriale française, le plus rapidement possible; c'est pourquoi l'information qu'il fournit aux libraires, à certains centres de documentation et bibliothèques est relativement succincte, mais vraiment d'actualité.

Nous pouvons remarquer qu'il se dégage deux grandes tendances dans les divers organismes visités.

D'une part, les bibliothèques à fonds encyclopédiques, conformes à la tradition de bibliothèques françaises. Même s'il apparaît, ces dernières années, des structures et des moyens nouveaux, leur rôle reste celui de mettre l'information à la portée de tous, dans le respect de l'esprit traditionnel de service public.

D'autre part les centres de documentation à fonds spécialisés dont la politique est liée à celle de l'organisme sous la tutelle duquel ils se trouvent.

Nous avons pu constater que la différence de politiques entraîne une production d'informations à contenus différents. Cela se répercute inévitablement sur la manière dont est traitée l'information et dont les documents sont indexés.

Je me suis personnellement intéressée aux bibliothèques généralistes (1). Parmi les organismes visités, l'étude qui suit va donc concerner la BN, la BPI, et la BM de la Part-Dieu. Ces bibliothèques ont mis en place un réseau national d'indexation. L'Ecole Polytechnique, bien que possédant un fond spécialisé pour un public restreint, a choisi de participer à ce réseau.

Enfin, cet étude est complétée par le point de vue du producteur du logiciel qui gère la liste de ce réseau, la Société ERLI.

(1) Les résultats de l'enquête concernant les bibliothèques spécialisées ont été développés dans le mémoire de Fabiola Rodriguez Lopez.

### 3. RESULTATS DE L'ENQUETE.

#### 3.1. De la liste Laval au réseau RAMEAU.

La Bibliothèque Nationale et la Bibliothèque Publique d'Information utilisent toutes les deux, pour indexer les documents, la liste d'autorité éditée par l'université de Laval.

La BPI, la première, a choisi ce vocabulaire contrôlé, dès 1974. La BN l'utilise depuis 1980.

Il s'agit de la traduction du répertoire de la bibliothèque du Congrès, qui a été entreprise dès 1946 par l'Université de Laval. La bibliothèque du Congrès l'a créé en 1914. Cette traduction est faite au fur et à mesure des besoins. Chaque nouvelle édition est donc une mise à jour de la précédente, et comprend des termes supplémentaires. La neuvième édition, actuellement la plus récente, date de 1983. La dixième devrait être publiée au début de l'année 1989.

Ce répertoire était, dans les années soixante dix, le seul vocabulaire contrôlé existant éprouvé par des grandes bibliothèques depuis un certain nombre d'années.

Il s'agit à la fois d'un vocabulaire, d'une grammaire et d'une méthode d'indexation. C'est donc beaucoup plus qu'une simple liste de mots-clés.

Ce système permet une indexation précoordonnée; l'indexeur va constituer des vedettes structurées de façon précise: chacune comprend une tête de vedette portant sur le sujet principal, puis des subdivisions qui affinent cette première notion. Les subdivisions peuvent porter sur différents aspects: thématique, géographique ou chronologique.

Exemple: personnes âgées--condition sociale--XVI<sup>e</sup> siècle.

On peut ainsi obtenir une indexation plus précise qu'avec des mots clés, qui permet d'aller assez loin dans les notions.

Le vocabulaire à proprement parler est encyclopédique. Il couvre à peu près tous les domaines du savoir, mais pas à un point de précision extrême. Il a une structure arborescente, avec des termes génériques (notés XX), des termes spécifiques ("voir"), des termes associés (notés VaXX) et des termes exclus (notés X). Il comprend donc un système de renvois identique à ceux des thésauri classiques.

La liste comprend à la fois des vedettes déjà toutes construites et aussi des termes qui peuvent être combinés entre eux pour en former d'autres. La structure de ces vedettes est très contrôlée, car les subdivisions sont précisément prévues par le système. Les subdivisions sujet sont de trois sortes: certaines sont affranchies et peuvent donc s'appliquer librement à différentes catégories de sujets. Par exemple, "condition sociale"

vient préciser les catégories de personnes, de pays et de villes. D'autres sont "esclaves", absolument indissociables d'une seule tête de vedette. Par exemple, "gare de triage" ne pourra suivre que le terme "chemin de fer". Une troisième catégorie de subdivisions sujet, intermédiaire entre les deux précédentes, s'applique seulement à une classe de sujets. Par exemple, "constellation" peut suivre seulement les noms des étoiles.

L'école Polytechnique utilise le répertoire Laval depuis 1976, sur les conseils de la BPI, et la bibliothèque municipale de la Part-Dieu depuis 1983.

Dans ces bibliothèques, les vedettes de leurs propres listes sont créées au fur et à mesure de leurs besoins. Lors du traitement des documents, si l'indexeur ne trouve pas de vedette dans Laval qui lui convienne, il peut traduire directement la liste du Congrès, ou si là non plus rien ne correspond à ce qu'il désire, il peut créer une nouvelle vedette, en respectant la syntaxe imposée par le système. Les chefs de services contrôlent par la suite ces créations, afin d'assurer une certaine cohérence de la liste. A la BPI, les indexeurs consultaient également l'ancienne liste de la BN, avant que celle-ci adopte elle aussi Laval.

Mais même si elles utilisent toutes le même répertoire, ces bibliothèques ont obtenu des listes matière assez différentes les unes des autres, après quelques années, notamment la BN et la bibliothèque de Beaubourg. Cette différence peut s'expliquer par le fait que l'indexation est par essence une activité subjective, très liée à la personnalité de l'indexeur et à la politique d'indexation choisie. D'autre part, les différentes éditions du

répertoire comportent des modifications importantes. Or la BPI a constitué sa liste à partir d'une version plus ancienne que la BN.

Afin d'uniformiser l'indexation en France, au moins au départ entre les deux grandes bibliothèques, la DBMIST a décidé de créer la cellule nationale de coordination de l'indexation matière, en Octobre 1983. Elle se trouve dans le service du CCN. Elle a pour but de favoriser un dialogue entre la BN et la BPI, de mettre en place une base de données et de former les bibliothécaires à Laval.

Dès le départ, l'idée était de créer un vocabulaire contrôlé, uniformisé au niveau national, à partir de Laval. Au début, la liste a reçu le nom de LAMECH (Liste d'autorité encyclopédique matière collective hiérarchisée). Mais le projet ne s'est concrétisé qu'à la fin de 1986, par la signature d'une convention officielle entre la DBMIST et la BN. On a alors assisté à la naissance de RAMEAU (Répertoire d'autorité matière alphabétique unifié). Cette convention confie à la BN la responsabilité de la cohérence intellectuelle de la liste. La cellule nationale s'occupe de la gestion du réseau des utilisateurs. Elles participent toutes les deux à la formation de ces derniers. La BPI, sans responsabilité officielle, a également participé au projet. Mais actuellement, elle continue à avoir sa propre liste, un peu différente de la liste nationale, car ses créations de vedettes ne sont pas contrôlées par la BN. (elle a tendance à créer des vedettes en prenant un peu plus de liberté par rapport à Laval).

La DBMIST a engagé les bibliothèques universitaires à participer au réseau, ce que certaines d'entre elles ont fait. Un certain nombre de bibliothèques municipales y ont également adhéré. Aujourd'hui, RAMEAU compte une cinquantaine de membres.

Le réseau des Centres Régionaux de Documentation Pédagogique, ainsi que le Centre National, pour une de ses sous-bases, sont membres de RAMEAU. Enfin, l'école Polytechnique en fait elle aussi partie.

La liste est alimentée au départ par la liste BN, enrichie par les créations des utilisateurs. Pour éviter une multiplication de créations anarchiques, les propositions sont contrôlées par la BN avant d'être validées et versées dans la liste. En principe, cette dernière ne refuse pas ces créations, mais a plutôt tendance à les modifier si elles mettent en péril la cohérence de la liste ou son caractère encyclopédique. La BN veille à ce qu'il n'apparaisse pas des excroissances beaucoup plus importantes dans certains domaines que dans les autres. La création des vedettes doit toujours suivre la même procédure: les utilisateurs commencent par consulter la liste nationale (pour l'instant encore essentiellement constituée de la liste BN), puis ils regardent la liste Laval et celle du Congrès. Ils peuvent ensuite créer un nouveau concept. En principe, lors de la mise à jour de la liste, la BN ne supprime pas de vedettes.

Cette liste RAMEAU existe sur microfiches, dont des jeux ont été distribués aux utilisateurs il y a un peu plus d'un an. Elle a été mise en ligne récemment par la société ERLI, à la demande de la DBNIST, sur le serveur du SUNIST. Elle est gérée par le logiciel Alexis, conçu au départ pour l'annuaire électronique du mini-

tel. L'accès en ligne est possible depuis le mois de Juin 1988, mais pas encore pour tous les adhérents.

La société ERLI, prestataire de services, a adapté son logiciel aux objectifs prévus conjointement par le ministère et la BN. Elle a créé un modèle conceptuel de données, qui a été révisé plusieurs fois, en vue des évolutions possibles.

Actuellement, Alexis permet la consultation de la liste, et la gestion de sa mise à jour, qui peut être faite par deux réseaux: directement par la BN, ou par les propositions de créations des adhérents contrôlées par la BN.

Ce logiciel gère des lexies, c'est-à-dire des chaînes de caractères, et des relations entre ces lexies.

Chaque lexie est affecté d'un type, ce qui signifie qu'on le fait entrer dans une catégorie. Par exemple, lorsqu'il gère un thésaurus, chaque mot sera affecté du type "descripteur/DE" ou du type "non descripteur/ND"

Les relations réciproques entre lexies permettent de choisir des stratégies, de construire différents types de modélisations de données. Le premier modèle retenu pour la gestion de RAMEAU était une représentation de la syntaxe de la liste. Ainsi, chaque lexie était affecté d'un type correspondant à une partie de la structure des vedettes de Laval. Les différents types possibles étaient les catégories "tête de vedette", ou bien "subdivision sujet" etc. Puis un deuxième modèle a été conçu. Il prend en compte la nature grammaticale des mots, en vue d'un passage en langage naturel, dans le futur. Par exemple, le lexie "ouvrage" est affecté du type "nom/NO". Chaque lexie ne peut être affecté que d'un seul type,

mais il existe également une autre sorte de relation qui permet de connaître la place du terme dans les vedettes.

Ce logiciel sait donc déstructurer les vedettes. Il est possible d'interroger la liste par n'importe quel terme. La réponse obtenue comprend toutes les vedettes dans lesquelles se trouve ce terme, quelque soit leur "poids". Une commande permet de trier ces réponses. On obtient alors toutes celles où le terme se situe en tête de vedette, puis celles où ce terme se trouve, toujours en tête de vedette, sous une forme un peu différente. Par exemple, le terme demandé peut se trouver dans une expression. Les réponses où il apparaît en sous-vedette s'affichent ensuite à l'écran.

### 3.2. Utilisation de ces systèmes.

La BN réalise le catalogage et l'indexation des documents imprimés directement en ligne, sur le système informatique Géac, en utilisant le format INTERMARC.

Pour les départements spécialisés, les documents de natures diverses (cartes et plans, estampes, photos et bientôt manuscrits) sont traités grâce à un logiciel "maison", OPALINE, implanté sur l'ordinateur BULL DPS7 de la BN. Le service informatique a adapté le format aux besoins spécifiques de chaque type de document.

La BPI, quant à elle, a jusqu'à présent utilisé le logiciel SCARABE, conçu spécialement pour elle. Mais elle envisage de changer de système au début de l'année prochaine.

A la BM de la Part-Dieu, le catalogage se fait encore manuellement, mais passera dans peu de temps sur Géac.

La bibliothèque de Polytechnique utilise le logiciel SIBIL, modifié pour ses propres besoins par le centre informatique de l'école.

Dans toutes ces bibliothèques, actuellement, les indexeurs disposent du répertoire manuel Laval et des microfiches de RAMEAU. Depuis le mois de Juin, certaines disposent aussi de la liste en ligne. Cette dernière est hiérarchisée et comporte tous les renvois. De plus, il existe de nombreuses notes d'application en ligne. La cellule nationale d'indexation a diffusé un guide de l'utilisateur.

Dans toutes ces bibliothèques, l'indexation est généralement très fine, très analytique. Dans tous les systèmes, la majorité des vedettes matières renvoient à une seule notice d'ouvrage. En revanche, les documents peuvent être indexés par deux ou trois vedettes. Seule l'école Polytechnique ne crée qu'une seule vedette par volume. Les indexeurs sont souvent influencés par les titres des ouvrages dans les disciplines scientifiques; en effet, les titres sont de plus en plus précis et explicites, face à la multiplication et la spécialisation des publications. Mais ce n'est pas le cas dans les autres domaines, où ils ne sont pas souvent significatifs.

En ce qui concerne le contenu de la liste nationale, on a pu constater une évolution. Au départ, Laval était utilisé sans modification. Puis on peut observer une réaction inverse, dans une période ultérieure, où les vedettes ont été créées de façon très

souple, un peu comme à la BPI. Actuellement, la tendance est plutôt de remonter à la source, d'aller chercher les vedettes directement dans la liste du Congrès et de les traduire, plutôt que de garder la traduction canadienne.

En 1988, les utilisateurs qui ont proposé le plus grand nombre de vedettes sont les bibliothèques du Musée de l'Homme et certaines BU, notamment celle de physique nucléaire de Grenoble. Il semble que la responsabilité de ces créations soit laissée aux BU CADIST.

L'école Polytechnique, qui a créé sa propre liste depuis 1976, dispose de la liste RAMEAU sur microfiches depuis un peu plus d'un an. Elle a donc entrepris de réviser sa liste, pour que les deux coïncident. Elle en est à la lettre "e", ce qui montre l'ampleur du travail à réaliser. Cela prouve que leur propre liste est assez éloignée de RAMEAU. Cette différence peut s'expliquer en partie par l'utilisation d'une version ancienne de Laval.

Tous les établissements qui ont opté pour le répertoire Laval l'utilisent pour indexer tous les documents imprimés et même souvent les documents sonores, visuels, que ce soit à la BPI, à la Part-Dieu ou à Polytechnique.

A la BN, les départements spécialisés envisagent tous de l'utiliser eux aussi dans un avenir proche. Les indexeurs des estampes et de la musique ont suivi des stages de formation au CCBT, le service de la BN chargé de l'indexation des imprimés et du contrôle de la liste nationale. Ceux du département des manuscrits enluminés envisagent de faire de même.

En revanche, le département des périodiques, qui indexait les titres avec Laval depuis 1983, ne le fait plus depuis le 1er Janvier 1988. Depuis cette date, les notices bibliographiques entrées dans la base BN-OPALE ne comporte plus qu'un indice CDU. Les bibliothécaires estiment que l'indexation matière n'apporte pas plus de renseignements que ceux contenus dans le titre, la collectivité auteur et l'indice CDU. D'après eux, pour un catalogue de titres de périodiques non dépouillés, l'indexation doit être beaucoup plus systématique qu'analytique. Ils ont donc élaboré un projet, intitulé PIRANAS (Procédé d'indexation et de recherche alphabétique et numérique appliqué aux séries). Il s'agit d'une table structurée qui permettra de faire des recherches grâce à des mots clés, et pas uniquement des indices numériques, à partir d'une indexation CDU. Cette table, plus restreinte que la CDU, est en cours d'élaboration. Elle devra permettre trois types de recherches, à partir des indices CDU: . une recherche présentée comme les menus minitel, avec trois indicateurs de plus en plus précis (le domaine, la rubrique, le descripteur);

. une recherche par mots clés, correspondant à des thèmes;

. une recherche à partir des indices de la CDU.

Pour l'instant, seule l'interrogation par l'indice numérique CDU existe.

Plusieurs BM, dont celle de la Part-Dieu, ont rencontré des problèmes pour l'indexation des ouvrages pour la jeunesse. La liste Laval comprend peu de termes dans ce domaine, et il ne s'agit pas d'une préoccupation prioritaire pour la BN.

Dans toutes les bibliothèques visitées, le public peut interroger les catalogues bibliographiques en ligne, au moins en ce qui concerne les publications et les livres français entrés depuis 1975, et les livres étrangers depuis 1983, à la BN.

Les catalogues de la BN et de la BPI sont interrogeables par le système Géac. La BN dispose d'un logiciel adapté à ses propres besoins, tandis que la BPI utilise le logiciel standard. Dans les deux cas, la liste matière en ligne n'est pas hiérarchisée, il s'agit d'une simple liste alphabétique de vedettes, avec des renvois associatifs et définitifs. Ces logiciels ne destructurent pas les vedettes. Pour interroger avec pertinence, il est donc préférable de connaître leur structure exacte. En revanche, il est possible d'interroger à partir des mots des titres. La BPI a mis à la disposition de son public une recherche par mots clés, avec le "et" booléen. Celui-ci peut également consulter son catalogue bibliographique sur CDROM. Ce dernier est couplé avec un logiciel d'interrogation qui peut destructurer les vedettes matières.

La BN vient également de terminer un prototype de CDROM, élaboré en collaboration avec la Bibliothèque Nationale anglaise, dans le cadre d'un projet européen. Cette concertation permettra aux différents pays de tous utiliser le même logiciel pour leurs catalogues sur CDROM. Ce CDROM comporte une partie des catalogues des deux bibliothèques nationales anglaises et française. Il va être distribué au début de l'année 1989 dans plusieurs bibliothèques françaises, afin d'avoir leur point de vue, puis il sera

disponible par abonnement, pour un prix d'environ huit mille francs par an, comprenant une mise à jour trimestrielle. Ce CDROM est ouvert, ce qui permet à ses utilisateurs de recopier les notices qu'ils désirent sur le disque dur de leur micro-ordinateur. Les recherches peuvent être faites par mots clés, c'est à-dire par mots significatifs, de n'importe quelle partie de la notice. Il est donc possible de faire une recherche par n'importe quel élément d'une vedette matière. L'utilisateur peut interroger sans risque de silence, sans connaître la structure du répertoire Laval. De plus, il est possible de faire une recherche multi-critères, en combinant plusieurs mots clés, avec les opérateurs booléens "et", "ou", "sauf". Enfin, les informations sont accessibles aussi bien en anglais qu'en français. Cet outil pourra être mis à la disposition du public de la BN dans avenir proche.

Qu'il soit manuel ou en ligne, le catalogue matière est très souvent consulté par les lecteurs, dans toutes les bibliothèques, même dans la salle des catalogues de la BN, où le public utilise également beaucoup les bibliographies spécialisées. Mais les bibliothécaires peuvent difficilement connaître le taux de réponses satisfaisantes obtenues.

Il n'y a qu'à l'école Polytechnique où le public fait la plupart des recherches à partir des auteurs ou des titres. Pourtant, il dispose d'un logiciel d'interrogation qui destructure les vedettes. Ce manque d'intérêt pour l'accès par sujets aux documents s'explique par la nature du public. Les lecteurs sont soit les élèves de l'école qui viennent principalement chercher les livres indiqués par leurs professeurs, soit des chercheurs qui connaissent très bien leur sujet et les références des ouvrages

désirés. De plus, la majorité" des documents appartiennent à des disciplines scientifiques, où les titres sont souvent très significatifs. Lorsqu'ils font une recherche par mots métiers, les lecteurs disposent de la liste en ligne, avec de nombreux renvois. Mais bien souvent, ils ne comprennent pas la syntaxe des vedettes et ils les utilisent comme de simples mots clés. Les bibliothécaires limitent donc de plus en plus le nombre de subdivisions des vedettes. D'autre part, ils proposent d'effectuer eux-même les recherches.

#### 4. CRITIQUES ET EVOLUTIONS POSSIBLES DES SYSTEMES.

La liste RAMEAU est donc issue de la liste Laval, qui est une traduction canadienne de la liste du Congrès. Les termes ne coïncident pas toujours avec le vocabulaire français, certains concepts n'ayant pas le même sens dans les deux pays. Il n'est pas évident non plus de traduire directement la liste américaine, car les termes choisis ne correspondent pas nécessairement à ceux utilisés spontanément par les français pour une recherche sur le même sujet. C'est pour cela que la liste RAMEAU a été créée, pour avoir un vocabulaire contrôlé véritablement adapté aux besoins des utilisateurs français.

D'autre part, les anglosaxons indexent de façon beaucoup plus systématique que les français, quitte à mettre plusieurs vedettes par ouvrage. La tradition française correspond à une indexation analytique très précise, parfois même trop pointue. Il y a donc un équilibre non évident à trouver entre les vedettes de

Laval, assez générales, et l'analyse excessivement précise des français.

A cette caractéristique s'ajoute le caractère encyclopédique de la liste, que RAMEAU entend conserver. Cela pose de gros problèmes aux bibliothèques qui possèdent des fonds spécialisés. Les BU qui participent au réseau national proposent un très grand nombre de vedettes nouvelles. C'est le cas, nous l'avons mentionné, de la BU de physique nucléaire de Grenoble. La BU de droit de Cujas a créé sa propre liste de vedettes matières, car elle estimait que les termes de Laval n'étaient ni assez précis, ni adaptés au droit français. Le nombre de termes du répertoire se rapportant au domaine juridique est en effet limité. De plus, le droit français est essentiellement légaliste, alors que le droit anglosaxon est jurisprudentiel. Les philosophies de ces deux systèmes juridiques sont éloignées. Certains concepts existent en droit français et pas en droit américain, et inversement. L'utilisation de Laval est donc spécialement difficile dans ce domaine. Mais la création d'une liste d'autorité matière ou d'un thésaurus demande beaucoup de temps et d'efforts et les BU françaises manquent crucialement de moyens.

L'école Polytechnique rencontre le même genre de problèmes pour son fond scientifique.

La BN, qui compte bien limiter toutes ces créations afin de conserver l'homogénéité de la liste dans tous les domaines, risque d'avoir à trancher de délicats problèmes.

d'oublier que l'indexation a une utilité pratique et ne doit pas rester un travail purement intellectuel. Certaines vedettes, très longues, avec beaucoup de subdivisions, sont très précises et constituent un travail scientifique important, mais souvent inutile au lecteur; et les vedettes trop longues ne sont pas adaptées à la configuration d'un écran informatique.

De toute façon, la liste RAMEAU est avant tout un outil pour les indexeurs, beaucoup plus qu'une aide à la recherche pour le grand public. La structure des vedettes est compliquée et nécessite une formation, qui a été prévue par la DBMIST et par la BN pour les bibliothécaires. Mais pour le public, ce n'est pas simple. Et le système Géac, que de plus en plus de bibliothèques adoptent, n'est pas très souple pour l'interrogation par vedettes matières, car il ne les destructure pas. Si on ne tape pas l'intitulé exacte d'une vedette, le risque de silence est grand. Or ces vedettes sont très éloignées du langage naturel. Si le système des renvois fonctionne parfaitement et si l'interrogation par mots matières est possible, le public est alors aidé.

Le logiciel documentaire Alexis présente plus de possibilités que les autres pour l'interrogation de la liste RAMEAU. Ses concepteurs envisagent, à terme, la possibilité de l'interroger en langage pseudonaturel. Mais la BN doit avant toute chose effectuer un gros travail de "nettoyage" de la liste, qui, telle qu'elle a été chargée, présente des incohérences. Les responsables de la société ERLI ont également songé à la possibilité d'offrir d'autres services, si les décideurs politiques le leur demandent. Notamment, ils envisagent de mettre en ligne le guide diffusé par la cellule nationale d'indexation. Le logiciel Alexis pourrait également être développé pour permettre la traduc-

Le but de l'indexation matière est de faciliter la recherche de l'information. Très utilisée, elle est indispensable dans les systèmes manuels et automatisés. Elle ne peut pas être remplacée par l'interrogation par les mots du titre. Cela pourrait éventuellement être envisagé dans une bibliothèque scientifique comme celle de Polytechnique. Dans ce domaine, nous l'avons vu, les mots du titre sont très significatifs. Mais une telle solution n'est pas concevable dans une bibliothèque encyclopédique. De même, on peut se demander si la décision du département des périodiques de la BN ne va pas entraîner un retour en arrière. La classification CDU ne joue pas exactement le même rôle qu'une indexation matière, même dans les cas où cette dernière est plutôt systématique. De plus, les concepts de la CDU sont souvent anciens, et peut-être difficilement applicables à certaines publications contemporaines.

L'indexation doit dans tous les cas s'adapter aux besoins spécifiques du public propre à chaque établissement. Les indexeurs doivent travailler dans cette optique. Dans la plupart des organismes, ils sont en contact direct avec le public. Ainsi, à la BPI, chaque bibliothécaire participe à toute la chaîne du livre, de l'acquisition à l'accueil des lecteurs, en passant par le catalogue et l'indexation. L'organisation du travail est la même à l'école Polytechnique. A la BM de la Part-Dieu, tout le monde participe aussi à l'accueil du public. Mais à la BN, le travail est réparti par tâches. Les personnes qui indexent ne sont pas celles qui cataloguent, ni celles qui renseignent le public. Les indexeurs ne connaissent pas les problèmes concrets qui se posent aux lecteurs, ni ce qu'ils pensent du catalogue matière ou de l'interrogation en ligne par ces vedettes matières. L'organisation de la BN a pour conséquence de couper les indexeurs des préoccupations du public. Dans ces conditions, il doit être facile

tion automatisée des vedettes. Dans le cadre d'un projet équivalent à celui du Pancatalogue, il serait alors possible de vendre aux anglosaxons des notices comprenant des vedettes matières. Le logiciel pourrait enfin évoluer vers un contrôle automatisé de l'indexation.

## 5. CONCLUSIONS.

La création d'une liste nationale d'indexation matière à partir du répertoire Laval pose de nombreux problèmes. La syntaxe semble compliquée et nécessite une formation, le vocabulaire est dans de nombreux cas mal adapté ou insuffisamment approfondi. Dans les petites BM, une simple liste de mots matières, par exemple la liste Blanc Montmayer, semble suffisante. Cette dernière ne permet pas de réaliser une indexation très pointue; mais les petites BM n'ont pas le même rôle à tenir que la BN, notamment, dont l'indexation doit être la plus scientifique possible.

Le contrôle de la cohérence de la liste représente un gros travail pour la BN, qui dispose de peu de moyens pour le réaliser. Même si elles se réfèrent à une liste d'autorité unique, les bibliothèques n'indexeront probablement pas leurs documents de façon identique.

Certains pensent qu'un tel réseau national ne devrait pas s'appuyer sur la liste Laval, mais sur un thésaurus. Il n'est pas évident de créer de toute pièce un nouveau vocabulaire contrôlé. Pour la BN, la simple gestion d'une liste nationale représente déjà un travail lourd. Pourtant, le centre de l'éducation nationale, a commencé l'élaboration d'un thésaurus encyclopédique. Pour le public, la recherche à partir d'un thésaurus est plus facile, dans un système informatisé, car il s'agit d'un langage postcoordonné, et les lecteurs peuvent utiliser le thésaurus comme une liste de mots clés. Mais si on arrive à obtenir une interrogation en langage pseudonaturel de la liste RAMEAU, comme les concepteurs d'Alexis le soutiennent, ce problème serait résolu. Et l'indexation est beaucoup plus riche avec ce répertoire, car les différents éléments des vedettes n'ont pas tous le même "poids". Chaque vedette est donc beaucoup plus précise que la combinaison de plusieurs mots clés.

Pour que cette liste puisse servir d'outil d'aide à la recherche pour le grand public, il faudrait prévoir un système d'interrogation à plusieurs niveaux, permettant d'utiliser la syntaxe de Laval, mais masquée par des interfaces. Le lecteur ne serait alors pas obligé de connaître la structure des vedettes. On pourrait par exemple imaginer une première interrogation par un mot clé, correspondant à la tête de vedette. Puis un mini menu proposerait à l'utilisateur un choix du type: "précisez la notion avec

1. un sujet

2. une date ou une époque

3. un nom géographique

Le choix tapé correspondrait à la première subdivision de la vedette. Il pourrait alors y avoir un deuxième menu, pour la subdivision suivante.

Une autre solution, à côté de l'utilisation du langage naturel, peut être trouvée dans les possibilités d'interrogation offertes par les logiciels couplés aux CDROM. Cette technologie permet au grand public d'effectuer des recherches par sujets sans connaissance de la structure des vedettes. Cela permet de réaliser une indexation très scientifique et très complète grâce à Laval, que le public peut retrouver très facilement, par simples mots clés. De plus, un tel système présente l'avantage d'être peu onéreux. Et si la possibilité de récupérer des notices aboutit à un catalogage partagé, seuls quelques catalogueurs formés à la liste RAMEAU réaliseront l'indexation matière. Les petites bibliothèques pourront alors bénéficier de ce travail qui nécessite beaucoup de temps et coûte cher. Mais l'information fournie par le CDROM n'est pas très récente, car la mise à jour suppose une nouvelle presse du disque. Cela peut être gênant pour une bibliothèque d'actualité. Cela pose également des problèmes si l'on met le catalogue sous cette forme à la disposition du public.

De plus, la conception de l'indexation doit être révisée. Les catalogueurs doivent s'adapter aux besoins du public, qui évoluent en fonction des outils mis à sa disposition. Avec les possibilités offertes par le CDROM, ce n'est plus la peine de multiplier les renvois. Actuellement, les indexeurs savent que Géac ne destructure pas les vedettes. Le choix d'un terme comme tête de vedette pouvant être arbitraire, ils font dans ce cas un renvoi de la subdivision sujet vers la tête de vedette.

Par exemple, lorsqu'ils créent la vedette: "chien--dressage", ils créent également un renvoi de "dressage de chien" vers cette vedette. Ces renvois ne sont absolument plus nécessaires avec le CDROM. Il permet donc d'alléger la gestion de l'indexation, qui est encore très lourde aujourd'hui. Les indexeurs doivent donc s'adapter aux nouvelles technologies. Mais cela n'est pas évident, quand on voit la vitesse à laquelle celles-ci évoluent, et d'autre part, le temps qu'il faut pour mettre au point un système d'indexation.

Il faudrait également expliquer au public qu'un système documentaire informatisé permet d'obtenir tout de suite une information très pointue. Les questions doivent donc être formulées le plus précisément possible. Or, en général, les utilisateurs font l'inverse. Lorsqu'ils interrogent un catalogue informatisé, ils ont souvent peur du silence. Par exemple, si leur recherche porte sur les fours à micro-ondes, ils ont tendance à taper le terme "cuisine", puis ils affinent leur recherche.

Finalement, le choix de ce vocabulaire contrôlé est très opportuniste. Il s'agissait du seul système à avoir déjà fait ses preuves dans des grandes bibliothèques.

Mais il fallait absolument résoudre le problème de l'indexation en France. C'est un des points cruciaux soulevés par la majorité des bibliothèques, qui ne savent pas exactement comment la réaliser. Lorsqu'elles s'automatisent, elles doivent avoir un vocabulaire strictement contrôlé, car les incohérences, anodines dans un système manuel, apparaissent tout de suite.

De plus, à l'heure actuelle, la coopération entre les bibliothèques se développe, notamment en ce qui concerne le traitement des documents. Si l'indexation matière était faite une seule fois par la BN, par exemple, puis réutilisée par les autres bibliothèques, ces dernières éviteraient de nombreux problèmes intellectuels et matériels. C'est pourquoi l'idée de créer un réseau national est vraiment intéressante. Le nombre important d'établissements qui y ont adhéré montre que le projet répond à un besoin. Les BU sont poussées par la DBMIST, mais pas les BM ni les autres membres du réseau. Les bibliothécaires de l'école Polytechnique nous ont exprimé leur volonté d'appartenir à ce réseau national, par esprit de solidarité, même si RAMEAU n'est pas le répertoire idéal pour indexer leur fonds scientifique. Les bibliothécaires que nous avons rencontrés ont l'air d'être vraiment motivés par ce projet. Avec le développement des nouvelles technologies, de nombreuses possibilités restent ouvertes. Mais en même temps, il est difficile, à l'heure actuelle, de savoir si le choix de ce vocabulaire contrôlé ne va pas être remis en cause par ces nouvelles technologies, dans avenir proche.

## EXTENSION DU FORMAT INTERMARC AUX

### MANUSCRITS A DECOR.

Le département des manuscrits occidentaux de la Bibliothèque Nationale compte plus de cent mille volumes, répartis par langues. Certaines collections anciennes sont toutefois demeurées autonomes. Ils couvrent plusieurs disciplines: l'histoire, la littérature, la philosophie, le droit, la théologie, les sciences et la généalogie.

Le département a pour mission de conserver ses collections, de les enrichir et de les communiquer aux chercheurs. Il publie des catalogues scientifiques de ses fonds.

La BN a passé une convention avec le CNRS, en application de laquelle les chercheurs de cet organisme travaillent en collaboration avec les conservateurs.

Le Centre de Recherche sur les Manuscrits Enluminés Médié-  
vaux (C.R.M.E.) est à l'origine d'un projet d'informatisation du  
catalogage. Environ dix mille manuscrits médiévaux sont enluminés.  
Ils constituent une source de documentation de grande valeur dans  
le domaine de l'histoire de l'art. De plus, le public vient faire  
des recherches d'images, pour des utilisations variées. Ce travail  
de recherche prend énormément de temps au personnel de ce service,  
alors que sa tâche prioritaire reste la publication des cata-  
logues. Le C.R.M.E. est donc spécialement intéressé par un projet  
d'informatisation, afin de disposer d'un outil d'aide pour la re-  
cherche des documents.

Le service informatique de la BN, auquel les responsables  
se sont adressés, a prévu de créer une base bibliographique de ce  
fonds avec OPALINE, comme pour les autres fonds spécialisés. Mais  
l'informatisation d'un service suppose l'utilisation d'un format  
de catalogage. Il a donc été nécessaire de rédiger une version du  
format intermarc spécialement développée pour les manuscrits à dé-  
cor.

### 1. Présentation d'OPALINE.

OPALINE est un logiciel de gestion documentaire qui a été  
conçu par le service informatique de la BN.

La base bibliographique des documents imprimés, BN-OPALE, étant déjà très importante (presqu'un million de notices, avec un accroissement annuel d'environ soixante-dix mille notices), la gestion des fonds spécialisés a nécessité la création d'une nouvelle base. Cette dernière ne se trouve pas sur l'ordinateur Géac, mais sur un ordinateur BULL DPS7.

Le département des Cartes et Plans a commencé à entrer ses notices dans cette base en Septembre 1987, et celui des estampes et de la photographie en Mars 1988, dans le cadre du projet de la BN sur la révolution. La Phonothèque Nationale s'est également informatisée, mais de façon complètement autonome, sans liaison avec le service informatique.

Plusieurs versions d'OPALINE ont été développées, afin de répondre le mieux possible aux problèmes des catalogueurs. Actuellement, la version Vo3, mise en place pour le catalogage des estampes, fonctionne également pour les cartes et plans.

Chaque département gère sa propre sous-base. Chaque sous-base comprend des fichiers bibliographiques, des fichiers d'autorité, des fichiers de liens et des tables. Ces dernières sont des listes d'autorité concernant les types de documents, les langues, les pays, les supports des documents, etc.

Chaque département gère ses propres autorités au sein de sa sous-base. Ces fichiers d'autorité portent sur les noms de personnes physiques, les collectivités, les noms géographiques et les mots matières. Mais les départements spécialisés ne sont pas des organismes autonomes. Ils font partie intégrante de BN. A terme, la BN voudrait offrir au public une base unique multi-médias dans laquelle se trouveront les toutes ses collections. C'est pourquoi, pour un utilisateur, la vision des autorités doit être unique. C'est ce qui se passe dans OPALINE. L'existence d'une telle passerelle est prévue entre OPALE et OPALINE dans l'avenir.

Le logiciel OPALINE permet de réaliser le catalogage bibliographique, la gestion des autorités et la recherche documentaire.

Le catalogage s'appuie sur un système descriptif qui est l'adaptation du format intermarc, utilisé pour le traitement des imprimés dans la base OPALE, aux documents de différentes natures des fonds spécialisés.

Le traitement se fait directement en ligne dans les départements des Cartes et Plans et des Estampes.

OPALINE permet d'enregistrer les entrées de façon très simple. Les magasiniers, lors de l'arrivée des documents dans le service (par le dépôt légal, les acquisitions, les dons ou les échanges), créent des pré-notices, qui comportent un nombre restreint de renseignements. La saisie est facilitée par le logiciel qui inscrit automatiquement sur l'écran les étiquettes du format correspondant aux champs à remplir, et leurs intitulés. Nous avons participé au catalogage aux Cartes et Plans pendant une semaine.

Il suffit donc d'un minimum de pratique pour réaliser ce "pré-catalogage". Ces pré-notices peuvent ensuite être rappelées au moment du catalogage et complétées par des informations supplémentaires sur le document.

Le catalogueur est lui aussi aidé par le système qui lui propose l'ensemble des champs, avec le numéro des étiquettes et leur intitulé. Il n'a pas besoin de mémoriser toutes les étiquettes, contrairement à ce qui se passe lors du catalogage avec le système Géac. Il existe aussi des contrôles automatiques. Par exemple, s'il ne remplit pas un champ obligatoire, il ne peut pas passer au champ suivant. S'il remplit un sous-champ avec un code qui ne correspond pas à la table à laquelle il est relié, le système ne l'acceptera pas. Il existe différents niveaux de notices, ce qui permet également des contrôles. Il est également possible de créer des notices par copie. Cette fonction fait gagner un temps énorme aux catalogueurs. Ils peuvent rappeler une notice déjà existante et en créer une nouvelle seulement en modifiant le contenu de certains champs, ou en ajoutant quelques informations à la notice initiale. Cela est très utile pour cataloguer des documents appartenant à une même collection. Par exemple, les cartes routières Michelin: bien souvent, seul le titre et le numéro change. Pour les cartes d'un même lieu, seule la date d'édition doit être modifiée d'une année sur l'autre, la plupart du temps. Enfin, les fichiers d'autorités peuvent être consultés en ligne, à tout moment, avec la possibilité de revenir à la notice traitée par une simple commande.

L'indexation des documents est réalisée, comme dans la base BN-OPALE, à partir du répertoire Laval. Des indexations multiples sont possibles. Pour une même notice, cinquante données de type différent peuvent être indexées. Ce nombre peut être porté à quatre vingt dix en cas de répétitivité.

OPALINE est aussi un logiciel de recherche documentaire, mis à la disposition du public. Celui-ci peut formuler ses questions en utilisant le "et" booléen, en croisant dix critères. Une fonction navigationnelle lui permet de passer du fichier bibliographique aux références ou à l'index matière à tout moment. Le service informatique vient de mettre au point une nouvelle version qui destructure les vedettes matières. L'interrogation est donc possible par n'importe quel terme, qu'il s'agisse de la tête de vedette ou d'une subdivision.

## 2. Elaboration d'un format pour les manuscrits.

Pour pouvoir informatiser le traitement des documents, il est indispensable d'avoir un format, c'est-à-dire une structuration précise des données des notices. Pour que le système informatique puisse reconnaître les éléments significatifs, il faut les mettre entre des bornes. Le format permet de mettre ces éléments dans des champs, eux-mêmes divisés en sous-champs. Ils sont tous introduits par des étiquettes qui renseignent sur la nature des éléments qu'ils contiennent.

Pour utiliser un format, il faut une très grande cohérence et des règles très précises de catalogage. Chaque type de document est catalogué de façon différente. Il fallait donc élaborer un format adapté aux spécificités du traitement des manuscrits enluminés. Ce travail a été réalisé en collaboration avec les spécialistes du département.

La BN utilise le format INTERMARC, qu'elle a conçu avec d'autres bibliothèques à partir du format d'échange UNIMARC. Tous les formats utilisés à la BN sont issus du format INTERMARC pour les monographies (M).

Le premier problème rencontré a été l'absence de normalisation du catalogage des manuscrits. Celui-ci suit des règles coutumières. L'élaboration d'un format a obligé les catalogueurs à trancher pour un certain nombre de points. Il a fallu décider quels éléments seraient mis dans les différentes zones. En partant des notices des catalogues, il faut déstructurer les éléments qui les composent, pour remplir les champs du format INTERMARC (M). Certains champs ont dû être ajoutés, car la description bibliographique des manuscrits comprend certains éléments qui n'existent pas dans les monographies. Par exemple, dans la zone 280 de la collation, une information concernant les différents supports des manuscrits a été ajoutée. A l'inverse, certains éléments n'existent pas ou n'ont aucun intérêt pour les manuscrits. Par exemple, la zone 245 du titre propre aurait pu être supprimée, car la majorité des manuscrits n'ont pas de titre.

Mais des éléments nouveaux ne sont ajoutés que s'ils sont indispensables, et s'ils ne peuvent absolument pas être mis dans une zone déjà existante. En effet, l'idéal serait d'avoir des formats qui se ressemblent le plus possible, car les différentes sous-bases d'OPALINE sont à terme destinées à former une seule base multi-médias avec OPALE. D'autre part, plus les formats seront proches les uns des autres, plus les échanges de notices seront faciles. Certaines concessions ont dû être faites, à l'issue de nombreuses réunions. Par exemple, la zone du titre propre a finalement été conservée.

Un deuxième point a consisté à déterminer les zones obligatoires, afin de prévoir des contrôles informatiques du catalogage.

Enfin, il a fallu également déterminer les sous-zones qui seraient indexées, c'est-à-dire celles à partir desquelles on pourra ensuite retrouver les documents.

La motivation essentielle du département des manuscrits à décor pour s'informatiser a été le désir d'obtenir un outil d'aide à la recherche des images. Il a donc voulu créer une base iconographique. Les notices devaient comprendre une description détaillée des images, ce qui a posé des problèmes particuliers. Par exemple, certains manuscrits sont composés de plusieurs textes, auxquels se rattachent plusieurs images. Or de tels problèmes n'ont évidemment pas été prévus dans le format des monographies. Il était nécessaire de créer un lien entre chaque texte et les images s'y rapportant, mais également d'établir un

lien entre tous ces textes d'un même manuscrit. On a pensé, au départ, créer une notice par texte, en mettant la description des images dans une ou plusieurs zones. Les différents textes formant une seule entité, parqu'appartenant à un seul manuscrit, auraient été attachés grâce à un lien de collection. Mais la description des images devant être très détaillée, elle était structurée comme une notice, avec une zone de titre, une zone auteur, etc. Il a donc été prévu de la mettre dans des sous-notices, correspondant au catalogage à niveau des monographies.

### CONCLUSION.

La réalisation d'un tel travail demande une très grande coordination entre les bibliothécaires et les informaticiens, et nécessite des concessions de la part des catalogueurs. Il s'agit d'un travail certainement plus compliqué pour les manuscrits que pour d'autres documents dont le catalogage est normalisé. Bien que le format soit à présent entré en machine, le traitement des documents ne sera pas immédiatement informatisé. Il est nécessaire de saisir au préalable un certain nombre de bordereaux, afin de s'assurer qu'il est bien adapté et que tous les cas possibles ont été prévus.

Un des intérêts de l'informatisation consiste à obliger les catalogueurs à normaliser le catalogage. La coopération entre les bibliothèques sera ainsi facilitée, au niveau national et international. Pour les manuscrits enluminés, ce problème est moins important dans la mesure où chaque document étant en principe unique, le catalogage partagé n'a pas d'intérêt.

Néanmoins, l'élaboration de ce format est attendue avec beaucoup d'intérêt par d'autres grandes bibliothèques étrangères, notamment la bibliothèque royale de Bruxelles. En effet, il s'agit du premier format adapté aux manuscrits.

ANNEXES

---

## QUESTIONNAIRE

### 1. HISTORIQUE

- a) A partir de quel vocabulaire contrôlé réalisez-vous l'indexation des documents ?
- b) Depuis quand utilisez-vous ce vocabulaire contrôlé, et pourquoi? Par qui a-t-il été créé, où, quand, et dans quel but ?
- c) Quelle est la structure de ce système. Quel est le contexte de chaque descripteur ; existe-t-il des renvois, des mots-associés, des termes génériques et spécifiques, des ramifications... ?

### 2. LES COLLECTIONS GERÉES

- d) Le vocabulaire que vous utilisez est-il également repris par d'autres personnes, ou bien sert-il à indexer et interroger d'autres bases de données ? Lesquelles ?

### 3. L'ELABORATION DU SYSTEME

- a) Comment l'analyse des documents est-elle faite ?
- b) Comment sont choisis les mots-clés ?
- c) Utilisez-vous fréquemment les mots du titre comme mots-clés ?
- d) Combien de mots-clés utilisez-vous pour indexer un document (au minimum, au maximum) ?
- e) Comment sont choisis les niveaux d'importance des mots-clés ?
- f) Existe-t-il une syntaxe particulière ou des relations entre les descripteurs ? Lesquelles ?
- g) Est-ce que toute la collection des documents est analysée et indexée ?
- h) Quel est le degré de satisfaction des utilisateurs face à l'indexation que vous réalisez ?

## 6. SATISFACTION DES UTILISATEURS

- a) Les utilisateurs peuvent-ils interroger la base par mots-matières ?
- b) Quelle proportion d'utilisateurs fait-elle ses recherches par mots-matières ?
- c) Pensez-vous que les termes utilisés spontanément par les interrogateurs dans leurs questions coïncident avec les mots utilisés par les indexeurs ?
- d) Ne pensez-vous pas que la recherche par les mots du titre des documents pourrait suffire aux lecteurs (surtout pour les documents des dernières années) ?
- e) Avez-vous des contacts avec les indexeurs pour leur communiquer les besoins des utilisateurs ?
- f) Pensez-vous que le résumé des documents faciliterait la recherche (en utilisant des mots du texte, des mots-clés ou du vocabulaire libre) ou bien serait-ce superflu ?

## 7. L'AVENIR DU SYSTEME

- a) Comment envisagez-vous l'évolution de votre système d'indexation sur le plan documentaire (votre système est-il viable à long terme, ou pensez-vous le modifier) ?
- b) Et sur le plan technique : pensez-vous informatiser votre système (indexer en ligne, disposer du vocabulaire contrôlé en ligne). Pensez-vous que l'utilisation de nouvelles technologies comme le CD-ROM soit envisageable ou souhaitable ?

TABEAU RECAPITULATIF DES MANUSCRITS ENLUMINES.

colonne 1: étiquette

- " 2: contenu
- " 3: premier indicateur
- " 4: deuxième indicateur
- " 5: codes
- " 6: information
- " 7: contrôle par table, fichier d'autorité...
- " 8: sous-zone à caractère obligatoire
- " 9: sous-zone à caractère répétitif

1	2	3	4	5	6	7	8	9
001	n° d'identification							
040	origine du ms.	/	/	\$d	pays décor	+	+	+
				\$a	pays copie	+		+
				\$x	<i>à pays du décor incertain</i>			+
				\$c	localisation			+
				\$p	nationalité artiste			+
041	langue	/	/	\$a	langue	+		+
				\$z	dialecte			+
044	date	/	/	\$l	siècle		+	+
				\$m	quart de siècle			+
				\$n	datation en clair		+	
045	décor	/	/	\$a	typologie	+	+	+
				\$n	estimation	+		+
090	cote	/	/	\$c	sigle bibliothèque			+
				\$d	sigle département			+
				\$a	cote du manuscrit	+		+
				\$j	mf ou f-s			
				\$h	n° de reprod. vd			
100	auteur principal (personne phys.)	/	/	\$3	famille			
	[zone répétitive dans son ensemble]		5	\$a	n° notice autorité	+		
				\$a	ou			
				\$m	élément d'entrée			
				\$d	élément rejeté			
				\$d	dates			
				\$e	qualificatif			
				\$u	numérotation			+
110	auteur principal (collectivité) voir zone 100							

1	2	3	4	5	6	7	8	9
141	titre uniforme [zone répétitive]	0 1	/	\$3 \$a \$i \$l \$q \$M	à valeur internat. à valeur nationale n° notice autorité ou titre titre de partie extr., abrégé... version <i>la note en clair</i>	+		+
143	titre de forme [zone répétitive]	/	/	\$a \$b \$e \$f \$g \$u \$m \$n \$j \$d \$l	titre de forme titre de partie nom pers. phys./mor. élément nom rejeté qualificatif numérotation nom lieu établissement année(s) date extr., recueil			+ + + + + + + +
144	titre de contenu [zone répétitive]	/	/	\$a	titre			
245	titre propre [zone répétitive]	1 2	/	\$a \$r	titre significatif titre diff. vedette titre reste du titre			
280	collation	/	/	\$g \$h \$a \$d	support principal supp. secondaire foliot./pagination dimensions			+
300	note générale	/	/	\$a	note			
307	reliure [zone répétitive]	/	/	\$a \$d \$e \$g \$h	origine date couvrure/matière iconographie note	+		+ + +
308	filigranes papier	/	/	\$t \$e	terminologie références bibl.	+		+ +
309	bibliographie	/	/	\$a	références bibl.			

1	2	3	4	5	6	7	8	9
310	comparaisons [zone répétitive] R	/	/	\$a \$b \$c \$d	ville institution... cote... date cv/cl			
330	note iconographique	/	/	\$a	note			
370	marques possession [zone répétitive] R	/	/	\$a \$g \$h \$i \$j \$k \$t \$z	personne/collect. armoiries monogramme/initiales devise emblème ex-libris info. topographique note			+ + + + + + +
600	indexation image [zone répétitive] R	0 1 6 7	x	\$3 \$7 \$a \$x \$y \$z	nom de personne collectivité nom commun nom géographique cf. vedette autorité n° notice autorité subdiv. non contrôl. tête de vedette subdiv. sujet subd. géographique subd. chronologique	+		+ + +
610	genre iconograph.	/	/	\$a	genre iconogr.	+		+
625	typologie textes [zone répétitive] R	6	x	\$a	nom commun cf. vedette autorité tête de vedette	+		
700	auteur secondaire (personne phys.) R  [zone répétitive]	/	/ 5	\$3 \$4  \$a \$m \$d \$e \$u \$4	famille n° notice autorité fonction ou élément d'entrée élément rejeté dates qualificatif numérotation fonction	+		+ +  + +
710	auteur secondaire R (collectivité) voir zone 700							
860	sous-notice [zone répétitive] R	/	/	\$p \$l	n° folio/page appel d'une zone			+ +

✓ Zone 100 Auteur principal personne physique

2

Premier indicateur de traitement: /

Deuxième indicateur de traitement: /

5 (famille)

Ces éléments de la notice bibliographique font l'objet de notices d'autorité

\$3 = numéro d'identification de la notice d'autorité établie à la personne

Le \$3 permet le transfert de tout ou partie de la zone vedette de la notice d'autorité.

En l'absence de notice d'autorité, l'information est donnée en clair comme suit:

\$a = élément d'entrée

\$m = élément du nom rejeté

\$d = dates

\$e = qualificatif

\$u = numérotation

\$u: le n° est donné en chiffres arabes

ex.: \$a Grégoire \$u09

\$4 code de fonction

✓ Zone 110 Auteur principal collectivité

Structurée et fonctionnant comme la zone 100

Indicateurs de traitement identiques à ceux des zones vedettes de la notice d'autorité correspondante

✓ Zone 280 Collation

Indicateurs de traitement: //

\$g = support principal  
\$h = support secondaire  
\$a = foliotation ou pagination  
\$d = dimensions

\$h: utilisé pour signaler les cas de supports multiples  
ex. papier encarté de parchemin: \$g papier \$h parchemin

\$a: -l'abréviation utilisée pour folio(s) est toujours f.  
et pour page(s) p.  
cette abréviation est séparée du nombre par un espace blanc  
ex.: 245<sub>A</sub>f.

-feuillet de garde du début et de la fin indiqués en chiffres romains,  
entre [] s'ils ne sont pas foliotés ou paginés  
éventuellement, indication entre (), après le chiffre, d'une pagination  
ou d'une foliotation alphabétique  
ex.: III f. (a-c)

-nombre de folios ou de pages du corps du manuscrit en chiffres arabes  
-chaque chiffre est suivi de l'indication f. ou p. si le système est  
différent pour chaque partie du ms.; sinon, l'indication est rejetée à  
la fin

ex.: IV p. # 230 f.  
[III] # 100 + [II] f.

-erreurs de pagination ou de foliotation indiquées à la fin entre ()  
ex.: V # 100 f. (foliotés 1-50, 61-110)

\$d: les indications sont données en mm., h x l

Ex.: 280 // \$g parchemin \$a II p. 58 f. \$d 280 x 160

Zone 308 Filigranes du papier

Indicateurs de traitement: //

\$t = terminologie  
\$e = références (et éventuellement note)

\$t: indication du ou des filigranes selon une terminologie fixe mais non close  
(voir Annexe)

\$e: réf. aux types recensés dans les ouvrages de base  
ex.: Briquet n° 000  
var. Heawood n° 000

✓ Zone 370 Marques de possession

Indicateurs de traitement: //

\$a = personne ou collectivité  
\$g = armoiries  
\$h = monogramme, initiales  
\$i = devise  
\$j = emblème  
\$k = ex-libris  
\$t = information topographique  
\$z = note (texte libre)

\$a: -rappel, par la fonction "copie", d'un nom de personne ou de collectivité déjà donné en 700 ou 710  
-ou mention: non identifié

\$g: indication, dans l'ordre du blasonnement mais sans élément de syntaxe, des:  
-partitions (coupé, tranché, écartelé...)  
-pièces (pal, fasce, sautoir, chevron...)  
-meubles

Les couleurs (métaux, émaux, fourrures) ne seront précisées que pour:  
-fasce (ou fascé)  
-bande (ou bandé)  
-croix  
-aigle  
-lion

Figure et couleur sont enchaînées par un trait d'union

ex.: aigle-sable  
croix-gueules

Le nombre de figures, quand il est supérieur à 1, est indiqué en chiffre arabe après le nom (toujours au singulier)

ex.: merlette 6  
aigle-sable 3

\$h: lorsque le monogramme n'est pas identifié, on donne les lettres qui le composent dans l'ordre alphabétique  
ex.: acdilu

\$i: la devise est transcrite telle quelle

\$k: -les mots abrégés sont complétés, en mettant entre () les lettres restituées  
ex.: Lib(er) s(ancti)...  
-les lettres ou mots disparus par suite de grattage, de déchirure... sont restitués entre []  
ex.: Liber sancti P[etri]

\$t: utilisé pour indiquer la présence du manuscrit en un lieu imprécis (soit sans nom de collectivité)  
ex.: Paris

\$z: utilisé pour toute précision complémentaire, par ex. pour une description complète d'armoiries non identifiées

BIBLIOTHEQUE NATIONALE

DOCUMENTS ET COLLECTIONS

01/08/88

11H25

0000 0000 0000 0 0 0 0 0000  
 0 0 0 0 0 0 0 0 00 0 0  
 0 0 0000 0000 0 0 0 00 000  
 0 0 0 0 0 0 0 0 0 0  
 0000 0 0 0 0000 0 0 0 0000

Département:  
 MANUSCRITS ENLUMINES (v03)DMS  
 LISTE DU FORMAT INTERMARC EN  
 BIBLIOGRAPHIQUE

ZONE	SZON	OBLIG	INDEX	REPET	TABLE	FORME	ISBD	INDIC	FOIDS	LIBEL	FICH ASSO
040	\$a		99	0	PAYS	T			000	payscopie	
	\$c			0					000	localisat.	
	\$d			0	PAYS	T			000	paysdecor	
	\$e			0	PAYS	T			000	paysdecor?	
041	\$p			0	PAYS	T			000	natartiste	
	\$a		99	0	LANG	T			000	langue	
044	\$z								000	dialecte	
	\$l	G	99	0					000	siècle	
	\$m			0					000	quartsiecl	
	\$n	G							000	datenclair	

090	\$a \$c \$d \$h \$j	G G G	99		FDND	T	000 000 000 000 000	cotemanusc siglebibl. sigledpt reprod.vd n°mf;f-s
100	\$a \$d \$e \$n \$u \$3		01	0 0		A D L	000 000 000 000 000 000	nom.pph. Z dates qualificat prénom numérotat. n°notauto.
110	\$a \$b \$c \$n \$p \$q \$3		02	0 0 0 0 0		A L	000 000 000 000 000 000 000	nom coll. X nom coll. lieu reste nom élt.rejeté autre qual n°notauto.
141	\$a \$i \$l \$n \$q \$3		99	0		L	000 000 000 000 000 000	tituniform tit.partie extrait... langues version n°notauto.
143	\$a \$b \$d \$e \$f \$g \$j \$l \$n \$n \$u		99	0 0 0 0 0 0 0 0 0 0 0		D D	000 000 000 000 000 000 000 000 000 000 000	tit.forme tit.partie date nom pers. prénom qualificat année extrait nom lieu établisse numérotat.
144	\$a		99				000	titcontenu
245	\$a \$r		99				000 000	titpropre restetitre
280	\$a \$d \$g \$h	G	99	0			000 000 000 000	foliot. dimensions supp.princ supp.sec.
300	\$a						000	note
307	\$a \$b \$d \$e \$g \$h		99	0 0 0 0	PAYS	T D Q	000 000 000 000 000 000	nom pays nom ville date rel. couvrure iconogrph. note
308	\$e \$t		99	0 0		Q	000 000	bib.filigr filigranes
309	\$a						000	ref.biblio

	\$c			D	000	cote	
	\$d				000	data cv/cl	
330	\$a				000	note icono	
370	\$a		0		000	possesseur	
	\$g	99	0		000	armoiries	
	\$h		0		000	monogramme	
	\$i	99	0		000	devise	
	\$j	99	0		000	emblème	
	\$k		0		000	ex-libris	
	\$t		0		000	info.topog	
	\$z				000	note	
600	\$a	11		M	000	ved.image	X
	\$x				000	subd.sujet	
	\$y	18	0	Y	000	subd.géo.	Y
	\$z	19		Z	000	subd.chron	Y
	\$Y	17	0	X	000	subd.sujet	Y
	\$3				000	n°notauto.	
	\$7		0		000	autre subd	
607	\$7			L	000	n°notauto.	
610	\$a	99	0	Q	000	genreicono	
625	\$a			Q	000	typo.texte	
700	\$a	01		A	000	nompersphy	Z
	\$d			D	000	dates	
	\$e		0		000	qualificat	
	\$m				000	prénom	
	\$u				000	numérotat.	
	\$3			L	000	n°notauto.	
	\$4		0	FNOC T	000	fonction	
710	\$a				000	noncollect	
	\$b		0		000	spous ved.	
	\$c		0		000	lieu	
	\$n		0		000	reste nom	
	\$p		0		000	elt.rejeté	
	\$q		0		000	autre qual	
	\$3				000	n°notauto.	
	\$4		0	FNOC T	000	fonction	
860	\$a				000	n°folio/p.	
	\$p				000	n°folio/p.	
	\$1		0		000	appel zone	

## S I G L E S

IMA	Institut du Monde Arabe
CDSH	Centre de Documentation Sciences Humaines
CNRS	Centre National de la Recherche Scientifique
BN	Bibliothèque Nationale
BPI	Bibliothèque Publique d'Information
BM	Bibliothèque Municipale
BU	Bibliothèque Universitaire
BS	Bibliothèque Spécialisée
DBMIST	Direction des Bibliothèques des Musées et de l'Information Scientifique et Technique
BIE	Bureau Internationa de l'Education
OCDE	Organisation de Développement Economique et Social
FRANCIS	Fichier de Recherches bibliographiques Automatisées sur les Nouveautés, la Communication et l'Information en Sciences humaines et sociales

