

Fichiers d'autorité : encore beaucoup de chemin à faire

par Jean-Paul Gaschignard

Bibliothèque du Cher

La recherche par sujets est celle qui est la plus spontanément utilisée par le public dans les catalogues informatisés. C'est aussi celle dont les performances sont les plus faibles.

Décrire les expériences menées aujourd'hui sur la constitution des fichiers d'autorité, et l'état de la récupération des notices d'autorité en France, ne donnerait

qu'un bilan assez maigre et, pour tout dire, assez décourageant. On peut heureusement espérer que la situation changera assez rapidement dans les prochaines années. Maintenant que la plupart des bibliothèques françaises sont informatisées, elles cherchent à améliorer la cohérence des accès à leurs catalogues. Elles mesurent les enjeux de ce travail, mais aussi ses contraintes.

Pourquoi des fichiers d'autorité ?

On pourrait, après tout, se contenter de recherches par mots de la notice, par des accès non contrôlés ou, pourquoi pas, par des accès plein texte au sommaire ou au texte lui-même. Cette solution est utilisée dans de nombreux systèmes documentaires informatisés.

Comme toujours dans la recherche documentaire informatisée, on retrouve deux stratégies extrêmes. La première est la stratégie du chaos, économique en temps de travail : on tolère un taux important d'erreurs, on ne contrôle pas les accès, l'utilisateur fera bien son chemin tout seul. La seconde est l'exhaustivité organisée, rêve impossible de générations de bibliothécaires et de documentalistes : car il est impossible de *tout* cataloguer impeccablement, avec des accès *parfaitement* contrôlés et hiérarchisés qui répondraient à l'avance à *toutes* les interrogations de tous les publics. C'est impossible aussi bien parce que le temps de travail est limité que pour des raisons plus fondamentales.

La stratégie du chaos est jouable, car elle demande peu de travail, mais laisse le lecteur se perdre dans le catalogue. L'exhaustivité organisée est impossible. La question, pour les bibliothécaires, pour les responsables de chaque établissement, est de savoir quel compromis ils choisissent entre ces deux extrêmes : jusqu'où vont leurs exigences de qualité ? pourquoi ? en fonction de quelles recherches documentaires, c'est-à-dire de quels lecteurs ?

Un usage encore balbutiant

La liste des défauts actuels des accès aux notices informatisées est longue.

– L'administration des catalogues prend énormément de temps et n'est donc jamais complètement faite : il reste toujours des auteurs répartis entre plusieurs accès (« Rabelais » et « Rabelais, François »), des titres erronés, etc.

– Les indexations par sujet sont peu compréhensibles. RAMEAU, qui se répand grâce à la récupération des notices bibliographiques de la BNF, est-il réellement adapté aux publics des bibliothèques de lecture publique ? Quels moyens donne-t-on au lecteur pour comprendre le langage documentaire utilisé, et donc pouvoir s'en servir ?

– Certains champs ne sont pas aujourd'hui gérés par des fichiers d'autorité, alors que les recherches sont assez fréquentes et le mériteraient. Par exemple, la normalisation ne reconnaît pas la notion d'autorités commerciales, ce que font BN-OPALINE et Électre sans que les systèmes locaux puis-

sent vraiment récupérer ces données de façon pertinente¹.

Cet article, qui ne traite que des bibliothèques de lecture publique, voudrait donc aborder trois questions :

– comment réduire le temps de travail dans la gestion des fichiers d'autorité ?

– comment pourrait-on faciliter la recherche par le public ?

– quels accès ou fichiers supplémentaires seraient souhaitables, avec quelles conséquences sur les formats ?

Il présente un certain nombre d'hypothèses de travail concernant le développement des fichiers d'autorité, et éventuellement les modifications qu'il faudrait envisager dans leur structure. Que se passera-t-il en fin de compte ? Les paris sont ouverts...

Réduction du temps de travail

Aujourd'hui, excepté la bibliothèque municipale de Fresnes, dont l'expérience n'est pas généralisable², les bibliothèques qui importent des notices d'Électre ou de la BNF doivent assurer elles-mêmes, notice par notice, la cohérence de leurs fichiers d'autorité. Travail de Sisyphe, travail de romain : immense et toujours à recommencer.

Le cédérom des autorités de la BNF ne comporte pas de fonction de récupération commode, c'est-à-dire automatisable, des notices. Il n'est donc réellement utilisable que pour la consultation. Cet usage très limité, et son prix, expliquent sans doute le faible nombre de ses clients.

L'association des utilisateurs de GEAC travaille dans la perspective d'une mise à jour automatique des autorités : quand l'interprétation de la BNF a changé, pouvoir importer automatiquement la nouvelle forme, et obtenir, sans intervention sur chaque notice, le remplacement de la forme précédente.

Techniquement, l'entreprise n'est pas impossible. Elle suppose pourtant que le fournisseur du logiciel ait développé

toute une série de fonctions spécifiques. Elle suppose aussi que la BNF organise la consultation de ses notices – et surtout celle du cédérom – en signalant les modifications, et en permettant certaines fonctions de déchargement. Le format de fourniture de notices par la BNF, d'après ses déclarations à la FULBI³, ne changera pas jusqu'à la mise en service du système informatique de Tolbiac : espérons que, dans la préparation des fonctions de récupération, la BNF saura travailler en étroite coopération avec les utilisateurs de ses données bibliographiques.

Question supplémentaire : Électre utilise aujourd'hui RAMEAU. Pour que cette mise à jour automatique soit possible avec des notices Électre, celles-ci devront comporter les numéros des autorités BNF – ou alors, les notices BNF devraient comporter les numéros de futures notices d'autorité créées par Électre.

Une nouvelle fois, il faut travailler à la complémentarité et à la coopération des deux principaux fournisseurs français de notices bibliographiques. Là encore, il faut espérer que les obstacles institutionnels ne seront pas insurmontables.

La mise à jour automatique des autorités répondra au problème, mais dans quelques années seulement. Son existence dépend de la pression du marché sur les fournisseurs de logiciels : en clair, elle n'existera que si beaucoup de bibliothèques la demandent dans leurs cahiers des charges ou par l'intermédiaire de leurs clubs d'utilisateurs. À chaque établissement de se déterminer.

Simplification de la recherche

Elle peut se faire par l'une des trois façons suivantes.

Indexations adaptées à la lecture publique

Première remarque : RAMEAU est inadapté aux publics de la lecture publique. Sa

1. Voir page 12.

2. La BM de Fresnes s'est abonnée, pour un coût annuel non négligeable, à la récupération de notices d'autorité de la BNF sur bande après requêtes ; chaque bande est retraitée par Opsy dans ses locaux.

3. Voir « Réunion Bibliothèque nationale de France/FULBI du 29 mai 1996, sur le déchargement des données disponibles sur les cédéroms de la BNF : compte rendu » : document interne à la FULBI et la BNF diffusé par les clubs d'utilisateurs de logiciels à leurs adhérents, ou par la BNF.

structure par vedette et sous-vedette est difficile à gérer par les systèmes informatiques⁴.

Comme le suggère Hervé Le Crosnier⁵, on progresserait en changeant radicalement les systèmes d'indexation par sujet, en construisant de nombreuses vedettes isolées (c'est-à-dire en abandonnant les constructions vedette + sous-vedettes). Est-ce réellement possible? Peut-on demander à la BNF de bouleverser tout son système d'indexation matière en fonction d'une réflexion prospective?

Si la BNF ne quitte pas RAMEAU, une autre solution serait d'*ajouter* une deuxième indexation. On touche ici à un paradoxe économique : aucun fournisseur de notices n'a intérêt à ajouter de nouvelles indexations, car le travail est important, et donc cher. Mais si le système d'indexation est inadapté, toutes les bibliothèques qui utilisent les notices en pâtissent. Logiquement, il faudrait que les bibliothèques utilisatrices paient le service supplémentaire qui leur est utile. Il faudrait pour cela une cogestion de tout le système de création et transmission des notices, impliquant ensemble les producteurs et les utilisateurs.

Est-il réaliste d'imaginer une telle coopération entre la BNF et les bibliothèques utilisatrices, ou entre Électre et les mêmes? Les contraintes institutionnelles contrarient les solutions techniques.

Pluralité des indexations

De toute façon, une seule indexation-matière est insuffisante dans un catalogue. Les publics sont nombreux, les attentes sont variées.

Va-t-on supprimer une indexation médicale dans une bibliothèque de médecine, sous prétexte qu'une indexation encyclopédique est plus adaptée à certains documents? ou l'inverse?

Les logiciels répandus dans les bibliothèques ne gèrent encore pour la plupart

qu'un index par type d'accès : mots-matières, indexation numérique... On peut espérer qu'ils permettront très prochainement plusieurs indexations simultanées : dans la même base, donner accès à un index sujet encyclopédique, un ou plusieurs index selon des listes d'autorité spécialisées par sujet, une indexation sujet pour les enfants, la Dewey et l'indexation des documents sonores...

Il n'y a pas là de difficulté majeure pour les fournisseurs de logiciels, mais il faut du temps de programmation. Cette pluralité d'index n'existera, là encore, que s'il y a une pression collective, globale, des bibliothèques, dans leurs cahiers des charges et les demandes des clubs d'utilisateurs.

Il y faudra aussi, sans doute, des adaptations des formats. UNIMARC, par exemple, utilise les zones 600 à 607 pour l'indexation-matière principale, en précisant en sous-zone 2 le code du système d'indexation employé. Il réserve par ailleurs une seule zone 610 pour une autre indexation. USMARC permet le libre usage des zones 690 à 699. En fait, il faudrait sans doute attribuer aux principaux systèmes d'indexation des zones ou séries de zones différentes, évitant toute ambiguïté, et permettant d'intégrer facilement dans la même base de notices de sources différentes. Avis au comité permanent de l'UNIMARC, avis à la communauté des bibliothécaires qui cherchent à faire avancer les questions de normalisation...

Maîtrise des langages d'indexation par les utilisateurs facilitée

Une grande différence entre le catalogue papier et le catalogue informatisé, c'est qu'il est beaucoup plus difficile de feuilleter un catalogue informatisé. L'utilisateur, à chaque écran, tend à supposer qu'il a terminé sa recherche. Il ne peut se faire une idée générale du langage d'indexation utilisé.

La solution réside sans doute dans l'usage de tableaux graphiques des indexations : sur un même écran apparaissent le terme sur lequel porte (ou peut porter) la recherche, ses termes associés, spécifiques, génériques, les termes les plus proches dans la recherche.

De tels écrans permettent une représentation synthétique de ce que les bibliothécaires ne maîtrisent qu'après de longues formations. Poussons un peu plus loin l'utopie : de cet écran montrant les différents choix les plus proches de notre terme, aller découvrir l'environnement du terme voisin, en déplaçant une surbrillance et en cliquant. De proche en proche, on peut ainsi monter et descendre dans l'arbre de l'indexation, en déplacer et en visualiser la carte.

Techniquement, l'augmentation phénoménale des puissances des ordinateurs, la conception même des architectures clients-serveur, rendront sans doute possibles, d'ici quelques années si ce n'est avant, de tels affichages graphiques.

Il existe bien aujourd'hui des interfaces graphiques (interfaces GUI) qui, à partir de la cote du document, le situent dans le plan de la bibliothèque, et même sur l'étagère où il se trouve grâce à une représentation en trois dimensions.

Cela suppose, une nouvelle fois, une évolution des formats de notices d'autorité. La notice devra donner, avec leurs numéros de référence, les termes génériques, spécifiques, associés. Il semble qu'INTERMARC autorités soit plus complet qu'UNIMARC autorités. Mais les formats ne sont pas immuables. UNIMARC sera amené à évoluer. Son comité permanent a du pain sur la planche.

Accès ou fichiers supplémentaires

Vers des autorités titre

Cette représentation graphique concerne en premier lieu l'affichage des indexations sujets. Elle peut aussi s'appliquer à d'autres index.

Elle pourrait entre autres servir à représenter les liens logiques entre les titres : titre d'une suite et titres séparés des volumes, liens entre un titre de volume et les titres d'œuvres différentes contenues dans ce volume... *Les Gens de Saint-Libéral* nous renverrait sur le même écran à *Les Palombes ne passeront plus* et *Des grives aux loups*. *Les Monumenta Germaniae historica*, aux multiples sous-séries imbriquées les unes dans les autres, trou-

4. Cette structure ne peut s'interpréter que comme une survivance des catalogues sur papier : pour des raisons matérielles simples (temps d'intercalation, taille des fichiers), on ne peut pas multiplier les accès dans un catalogue sur fiches ; ceux-ci ne permettent pas de combiner les recherches. La structure en vedettes et sous-vedettes permet de surmonter ces contraintes – mais l'informatique les fait disparaître.

5. Voir « Catalogues du futur, futur des catalogues », Journée d'étude. Actes, dans ce numéro p. 155.

veraient enfin un affichage facilement compréhensible (!). Et pourquoi ne pas utiliser de telles solutions pour visualiser les enchaînements de titres des périodiques (changements de titres, fusions, scissions, suppléments, etc.)?

On pourrait avoir une représentation simple des titres des morceaux contenus dans un disque, et de là aller voir les autres disques contenant l'un des morceaux...

Problème : il n'existe pas aujourd'hui de fichier d'autorité des titres. Mais il suffirait, en somme, de compléter et généraliser des fichiers aujourd'hui réservés, dans certains systèmes, aux titres uniformes. Là encore, cette fonction n'existera que si de nombreuses bibliothèques en font la demande. On peut également souhaiter que les formats évoluent en fonction de réflexions prospectives, plutôt qu'après que l'incohérence se soit installée dans les catalogues.

Des autorités titre aux autorités auteur-titre

La question de l'affichage des œuvres contenues dans un même document, et la gestion des liens avec le titre général qui les contient, ou avec d'autres titres à leur tour contenus dans les œuvres contenues, n'est pas entièrement résolue avec la généralisation des autorités titres. Carmen-Mozart sévit toujours⁶.

Le problème se complique encore si l'on réalise qu'un même titre peut être contenu dans un volume plus vaste et isolé dans un volume séparé, voire à son tour éclaté en plusieurs volumes différents (par exemple dans une édition en gros caractères).

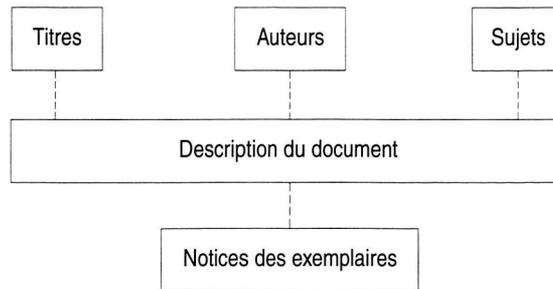
L'UNIMARC théorique (tel qu'il est défini par l'IFLA) ne permet jusqu'ici d'éviter Carmen-Mozart qu'en créant une sous-notice complète, ce que bien peu de systèmes savent faire. USMARC prévoit des entrées auteur-titre, mais semble-t-il (je ne connais pas très bien USMARC et apprécie-

6. Syndrome selon lequel une requête auteur-titre aboutit à un résultat même si l'auteur n'a pas créé ce titre, mais un autre contenu dans le même document. Pour plus de détails sur le syndrome Carmen-Mozart, voir D. Lahary, J.-P. Gaschignard, « L'affaire des liens », *Bulletin d'informations de l'ABF*, n° 163, 2^e trimestre 1994, p. 56-60.

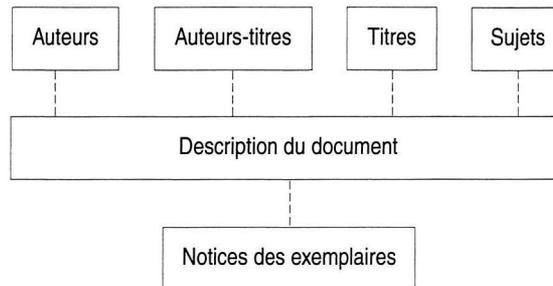
rais de recevoir des précisions sur ce point), sans renvoyer à des fichiers d'autorité, en permettant uniquement la constitution d'un index spécialisé.

Il faudrait pour bien faire créer un nouveau type de notice d'autorité, renvoyant à la fois aux autorités titre et aux autorités auteur pour établir les liens entre elles.

Du schéma actuel :



on passerait à des relations de ce type :



La notice d'autorité auteur-titre devrait contenir à la fois les numéros des autorités auteurs concernées, le numéro de l'autorité titre concernée, et les liens avec les notices bibliographiques.

Seule une solution de ce type permettrait à la fois de gérer d'une manière unifiée les différentes formes des noms d'auteur, les relations entre titres, et d'afficher de manière synthétique et facilement compréhensible les choix possibles pour trouver l'œuvre. Il s'agit en fait de combiner le catalogage du document et un catalogage des œuvres. Quand le lecteur cherche un titre, il cherche une œuvre avant de chercher un livre ou un document sonore.

Une variante technique serait de faire remplir à un seul fichier d'autorité les rôles d'autorité titre et d'autorité auteur-titre. La notice d'autorité de l'œuvre comprendrait alors les liens avec les auteurs, les liens avec les notices bibliographiques, les relations avec les titres rejetés, inclus, titre de l'ensemble de niveau supérieur...

On pourrait penser que ces autorités auteur-titre feraient double emploi avec les fonctions de catalogage analytique (création d'une notice-mère liée à des sous-notices), que proposent de plus en plus de logiciels. En fait, c'est poser en même temps deux questions différentes : – les autorités auteur-titre s'imposent si l'on cherche à faciliter la recherche des œuvres ;

– le lien avec ces autorités ne remplace pas la création de sous-notices, très utiles dans certains cas (dépouillements de périodiques, etc.) ; il devient une alternative plus simple, permettant de créer des accès aux œuvres contenues sans saisir d'autre information.

De toute façon, nous sommes là dans les hypothèses. Il faudrait pour y parvenir, là encore, une demande assez collective des bibliothèques, et un gros travail de définition technique et de normalisation, à mener en concertation avec les fournisseurs de logiciels. Cette hypothèse est-elle réaliste ? Si elle correspond à des besoins réels des lecteurs, une solution de ce genre fera, tôt ou tard, son chemin.

Vers des autorités commerciales

La recherche par éditeurs est très utile pour les bibliothécaires, par exemple pour préparer les commandes, et encore plus aux disothécaires. La création de ces fichiers paraît elle aussi inéluctable.

Pour que la notion d'autorité commerciale, que certains logiciels proposent, soit un outil et non une complication supplémentaire, deux conditions doivent être remplies :

- on doit pouvoir rattacher une notice bibliographique à plusieurs autorités commerciales, en cas de coédition ou de types d'autorité différents (comme dans BN-OPALINE) ;
- les fournisseurs d'information bibliographique doivent fournir des formes fixes

pour les éditeurs, ce que fait Électre mais non la BNF avec ses notices de livres⁷.

Pour l'instant, les règles de catalogage, respectées en cela par la plupart des formats MARC, ne traitent l'adresse bibliographique que comme une donnée descriptive non normalisée. Elles ne prévoient donc pas de liens avec des autorités.

Conclusion

Il y a donc du chemin à faire. Techniquement, aucune des innovations souhaitées ici n'est vraiment révolutionnaire. Il ne s'agit que de l'extension de fonctions utilisées dans certains systèmes existants. À part l'affichage graphique des liens entre

7. La BNF utilise des autorités commerciales dans son cédérom de documents sonores, audiovisuels et multimédia, mais pas dans son cédérom BNF qui contient des notices de livres français entrés depuis 1970. Peut-on ajouter que ces deux emplois d'un même sigle, par le même établissement, prêtent à confusion ?

autorités, tout cela est même réalisable sans Windows, sur des écrans noir et blanc en 80 colonnes très traditionnels, et déjà bien démodés aujourd'hui.

Ces innovations, moins spectaculaires que d'autres sujets en débat, ne se feront que s'il y a une demande collective des bibliothèques, par deux moyens : les cahiers des charges et l'action des clubs d'utilisateurs de logiciels.

La rédaction des cahiers des charges dépend de la formation de ceux qui les rédigent. Les responsables informatiques des bibliothèques se forment aujourd'hui sur le tas, en confrontant pas à pas leurs expériences. Intensifier et organiser ces débats, améliorer les outils de formation dans ce domaine, traiter non seulement des évolutions à long terme, mais de l'analyse des difficultés concrètes et immédiates, représentent aujourd'hui un enjeu majeur pour la communauté des bibliothèques.

Les clubs d'utilisateurs de logiciels se structurent progressivement en France.

Certains se sont organisés en groupes par module, voire en sous-groupes de travail, et recueillent des cotisations importantes. L'activité d'autres se limite à des assemblées générales, et quelques-uns ne disposent d'aucun moyen qui leur soit propre. Leur action est pourtant essentielle. En permettant l'échange des expériences, en faisant pression sur les fournisseurs de logiciels pour obtenir des améliorations, en débattant, à partir de leurs besoins immédiats, des perspectives de l'informatisation des bibliothèques, ils peuvent obtenir des fonctions adaptées aux bibliothèques et à leurs lecteurs.

Il est permis de penser que l'adaptation des fonctions est plus importante que les affichages des écrans. Un compte rendu approprié peut faire gagner des centaines d'heures de travail par an à une bibliothèque, et cela vaut le coup de se battre des années pour l'obtenir. Une recherche documentaire perfectionnée, facile à utiliser par les lecteurs, est tout sauf un enjeu négligeable.