



IFLA 96
62nd IFLA General conference
(25-31 August 1996)

Division number : Division du
contrôle bibliographique

Professional Group : Section de
bibliographie

THEME : CONTROLE DES
DONNEES D'AUTORITE

Les normes internationales sur le contrôle des données d'autorité : coûts et profits

par

Alan Danskin

Responsable du contrôle des autorités
The British Library

Introduction

Coïncidence le thème de la Section de catalogage cette année est "faire des économies en partageant les ressources". De grands pas ont déjà été faits en ce sens par l'échange de notices bibliographiques et c'est ce qui a suscité l'intérêt pour l'échange des données d'autorité. En dépit de quelques réalisations de coopération nationale, dont le Programme Name Authorities Co-Operative (NACO = Coopération pour les autorités auteurs) est un brillant exemple, peu de données d'autorité sont échangées à l'échelon international. Contrastant avec le succès que l'IFLA a rencontré en encourageant la création de bibliographies nationales et l'échange de données bibliographiques, l'effort consacré à appliquer ces principes aux autorités n'a pas produit les résultats escomptés. Après avoir examiné quels bénéfices on peut attendre de l'échange des données d'autorité, j'étudierai jusqu'où on peut aller dans leur réalisation.

Les économies induites par la coopération

Chaque bibliothèque peut tirer profit du partage des données d'autorité de 3 façons :

- comme source d'information pour identifier les vedettes d'autorité.
- en répartissant les coûts de création et de gestion des notices d'autorité grâce à l'accès à un fichier d'autorité partagé.
- en rendant encore plus performant l'échange de notices bibliographiques.

Malgré tout l'intérêt d'un échange international des données d'autorité, on a fait peu de progrès dans ce domaine. Il est beaucoup plus facile d'identifier les économies induites par le partage des ressources que d'en profiter. On n'a rien pour rien.

Le contrôle bibliographique universel et les données d'autorité

L'IFLA a contribué à développer les normes sur la forme et l'identification des données d'autorité avec les *GARE*¹ le format *UNIMARC(A)*² et la publication de Françoise Bourdon : *International Co-operation in the field of Authority data*³. On a eu tendance à appliquer aux données d'autorité les principes du contrôle bibliographique universel définis pour les données bibliographiques. Je crois que ces principes conviennent mal à ce domaine et leur mise en oeuvre sera d'un coût prohibitif.

Ces principes sont les suivants :

- une agence bibliographique nationale (ABN) est responsable de ses auteurs nationaux
- les agences bibliographiques nationales doivent établir les données d'autorité de référence pour leurs auteurs nationaux
- une agence bibliographique nationale doit diffuser les données de référence établies pour ses auteurs nationaux
- chaque vedette d'autorité doit être identifiée par un numéro unique, le Numéro international normalisé de données d'autorité (ISADN = International Standard Authority Data Number).

On a déjà beaucoup écrit sur les difficultés d'imposer une forme particulière de vedette à toutes les bibliothèques et je n'ai pas l'intention de reprendre ici le débat : il suffit de rappeler que Barbara Tillett⁴ et d'autres ont recommandé un nouveau type de notice pour contrôler les accès, qui servirait surtout à rassembler les variantes de formes d'un nom donné. Selon ce modèle l'ISADN joue le rôle d'identifiant unique, et donc il n'est pas nécessaire que la vedette remplisse elle-même ce rôle.

Qu'advient-il de toutes les notices d'autorité qui existent déjà ? Les agences bibliographiques nationales ont investi des millions de livres sterling pendant des années pour mettre sur pied des fichiers d'autorité auteurs (voir le tableau joint). Tous les fichiers d'autorité nationaux doivent-ils être traités de la même façon et se voir attribuer un unique ISADN en fonction de leur nationalité ?

Il est fondamental pour le contrôle bibliographique universel (CBU) que la nationalité des auteurs (et donc l'agence bibliographique nationale responsable de l'établissement

de leur nom) puisse être facilement identifiée. Cependant, un rapide examen des fichiers d'autorité nationaux montre que la plupart des agences bibliographiques nationales ne font pas cette distinction (voir tableau). Ainsi ni la British Library ni la Library of Congress n'indiquent la langue ou la nationalité dans leurs données d'autorité ; la Biblioteca Nacional de Madrid et l'Instituto da Biblioteca Nacional e do Livro (Portugal) indiquent la langue mais pas la nationalité ; l'université d'Helsinki et la Bibliothèque Nationale de France font exception en indiquant à la fois la langue et la nationalité. Même s'il était possible d'ajouter rétrospectivement la mention de nationalité, les discussions sur quelle nationalité attribuer aux auteurs personnes physiques persisteraient. Comme le dit Samuel Johnson : "C'est de ses auteurs qu'un peuple tire sa plus grande fierté". Mais il n'est pas raisonnable d'espérer des autres agences qu'elles retardent leur catalogage en attendant que les noms aient été établis par l'agence bibliographique nationale concernée, et ce sans avoir aucune garantie que toutes les agences puissent rendre largement disponible leur fichier d'autorité sous une forme adéquate et à un coût acceptable.

Etablir une unique forme préférentielle, si tant est que cela soit possible, serait insuffisant en soi : il faut aussi penser à la gestion des autorités dans le temps. L'instabilité des données d'autorité est une des caractéristiques qui les distinguent des notices bibliographiques. En général, une fois créée, une notice bibliographique évolue peu ; de plus, tout changement est normalement effectué au niveau local et est limité à une seule notice dans un fonds unique. Cette autonomie de leur contenu rend les notices bibliographiques faciles à échanger. Quel que soit notre désir de voir les données d'autorité présenter la même stabilité, nous devons admettre qu'elles doivent faire l'objet d'une mise à jour chaque fois qu'il faut ajouter des formes rejetées ou des informations sur les sources, lorsque les auteurs changent de nom ou lorsqu'il faut résoudre une homonymie par l'ajout d'un qualificatif de date. Les modifications concernant une notice d'autorité engendrent une dynamique : elles ne se limitent pas à une seule notice bibliographique, ni même à un seul catalogue; elles peuvent engendrer des modifications dans d'autres notices d'autorité ou la création de nouvelles notices d'autorité. Si on facilite l'échange international des données bibliographiques en développant la cohérence des données d'autorité, il est indispensable que les utilisateurs aient accès aux données les plus récentes.

Quel format doit-on utiliser pour échanger des données d'autorité ? L'IFLA a élaboré l'UNIMARC(A), mais la majeure partie des données d'autorité qui existent au monde sont en USMARC(A) ou dans un format dérivé de l'USMARC(A). A l'origine le format UNIMARC(A) était conçu comme un format pivot mais, c'est bien connu, les données sont altérées ou perdues au cours de la conversion et par conséquent l'intégrité des données transférées entre des systèmes utilisant des formats différents n'est pas garantie. Ceci n'a pas été un obstacle important pour l'échange de données bibliographiques, parce que celui-ci ne se fait généralement que dans un sens et parce qu'un certain niveau de dégradation des données est acceptable. Une altération des données d'autorité n'est pas acceptable parce que des points d'accès erronés pourraient être générés et parce que les données d'autorité étant sujettes au changement il pourrait y avoir une érosion progressive des données au fur et à mesure des transferts.

Ce sont là de grands obstacles à la réalisation des objectifs du CBU et les ressources nécessaires pour les surmonter sont immenses. Le profit que l'on pourrait tirer d'un

investissement si incertain est négligeable, et même il en résulterait probablement une baisse de la qualité des accès dans n'importe quel catalogue donné.

On peut réellement tirer profit de l'échange international des données d'autorité, mais on y réussira mieux en concentrant les ressources là où il y a le plus à gagner. Il existe déjà des projets qui proposent des modèles et peut-être les bases d'une meilleure structure de coopération. Deux approches complémentaires peuvent être envisagées : la première est de développer une coopération plus étroite entre les bibliothèques travaillant au sein d'une même culture "catalographique", ce qui englobe pour moi la langue de catalogage, les normes de catalogage et l'industrie éditoriale. Ces bibliothèques peuvent essayer de réaliser de réelles économies en développant un fichier d'autorité partagé et accroître par là l'efficacité avec laquelle les données bibliographiques sont échangées. La seconde approche consiste à rendre les données d'autorité aussi disponibles que possible grâce à des accès à distance. Cela réduira la multiplication des recherches pour identifier ou lier des vedettes et devrait rendre les agences bibliographiques nationales appartenant à différentes cultures catalographiques capables d'harmoniser le plus possible la forme de leurs vedettes dans les limites imposées par les normes nationales.

Première Etude de cas : le Fichier d'autorité anglo-américain

Le Fichier d'autorité anglo-américain (= The Anglo-American Authority File = AAAF) est une entreprise ambitieuse de la Library of Congress et de la British Library visant à élaborer un unique fichier d'autorité partagé pour remplacer les fichiers existants, la British Library Name Authority List (BLNAL) et le United States National Authority File (USNAF). Le projet s'articule en 3 phases.

La première préoccupation de l'AAAF a été d'étudier les divergences dans les pratiques de catalogage et les créations de notices entre la British Library et la Library of Congress. On a découvert un bon nombre de divergences, mais au cours de négociations sur les points de détail en 1993-1995, elles ont été soit résolues soit, dans quelques cas, exclues de l'AAAF. Je regrette d'avoir à dire que nous n'avons pas pu nous mettre d'accord sur la façon de romaniser les caractères chinois. La British Library utilise la transcription pinyin depuis 1958⁵, mais la Library of Congress utilise encore le système Wade-Giles. Les vedettes romanisées ont été exclues en attendant que la Library of Congress décide d'adopter le Pinyin. L'autre grande catégorie d'exclusions concerne les titres uniformes pour lesquels aucune conversion adéquate ne pouvait être faite entre les formats MARC. La plus grosse difficulté qu'a rencontrée le Projet AAAF jusqu'ici a été de définir une méthodologie permettant de concilier les particularités propres à chacun des formats MARC sans altérer les données à échanger. L'annonce faite en janvier de cette année selon laquelle les formats USMARC, CANMARC et UKMARC s'aligneraient sur un format unique d'ici 1999 traduit bien qu'il y a des limites à la conversion d'un format.

La première phase de réalisation, achevée au début de cette année, a été de charger une copie du fichier USNAF dans le système de catalogage de la British Library. Pour sauvegarder l'intégrité des données, USNAF a été chargé en USMARC. La mise à jour de cette copie à la British Library est assurée par un transfert de nuit, à partir du fichier maître de la Library of Congress, des notices créées et mises à jour. On peut interroger simultanément les fichiers USNAF et BLNAL et les catalogueurs à la

British Library utilisent déjà le fichier USNAF comme un fichier-ressource. Quand on a besoin d'une vedette pour laquelle aucune entrée n'existe à la British Library, la notice peut être copiée du fichier USNAF dans le fichier BLNAL au lieu de créer une nouvelle notice d'autorité. On est sûr ainsi que la British Library ne refait pas le travail déjà réalisé par la Library of Congress. La Library of Congress bénéficie du transfert des nouvelles données d'autorité créées par la British Library, données qui sont en progression depuis 1994. Cependant à long terme cet échange de données a un intérêt limité, puisque aucune bibliothèque ne peut assurer le suivi des modifications qui interviennent sur les données qu'elle a fournies. Au fil du temps il est inévitable, à moins que les mises à jour puissent être fournies automatiquement, que les formes des vedettes divergent.

Dans la Phase 2, la mise au point d'un logiciel de conversion MARC sophistiqué rendra la British Library capable de glaner des notices d'autorité entre USMARC et UKMARC. Jusqu'à ce que l'harmonisation des formats MARC soit réalisée, seules les notices qui peuvent être converties sans que leur contenu soit modifié seront admises dans le fichier AAAF. Il sera possible de créer une notice d'autorité à la British Library en UKMARC, de la convertir en USMARC et de l'envoyer directement dans le fichier maître USNAF de la Library of Congress. Un catalogueur travaillant à la Library of Congress, ou dans l'une des quelque 200 autres bibliothèques du réseau NACO, pourra l'utiliser, et si nécessaire la mettre à jour par ajout ou correction des données. Si la notice est modifiée de quelque façon que ce soit elle sera de nouveau fournie dans les 24 heures à la British Library et aux autres bibliothèques du réseau NACO. La Phase 2 est en cours et devrait s'achever cette année.

La troisième Phase du Programme AAAF sera l'harmonisation rétrospective du fichier d'autorité de la British Library (BLNAL) avec le fichier USNAF. Les formes de vedettes qui doublonnent seront éliminées, ce qui constituera un travail d'envergure. Le fichier BLNAL compte quelque 700.000 formes d'autorité auteurs. Il est raisonnable de penser qu'environ la moitié d'entre elles correspondra à des entrées existantes dans le fichier USNAF et que parmi celles-ci la moitié correspondra à une variante de forme. Cependant au minimum 175.000 vedettes d'autorité auteurs devront être corrigées. On a accepté le principe général selon lequel on préférera la forme du fichier USNAF. La Phase 3 ne sera pas réalisée avant 1999, d'ici là la British Library aura un fichier d'autorité lié à ses catalogues.

J'ai décrit en détail le programme AAAF parce qu'il illustre la somme de temps et d'énergie nécessaire pour effectuer l'intégration de données d'autorité créées par différentes agences bibliographiques nationales. Un tel investissement doit être justifié. La British Library crée 60.000 notices d'autorité auteurs par an, et une étude laisse entendre que 44%⁶ peuvent être des doublons de notices créées par la Library of Congress. Notons que les doublons de notices sont différents des doublons de vedettes. La British Library dérive environ 50% de son catalogage en langue anglaise de notices de la Library of Congress mais, puisqu'elle n'utilise pas le même fichier d'autorité que la Library of Congress, chaque notice bibliographique doit être mise en adéquation avec le fichier BLNAL. Cependant on n'a pas encore dérivé tout le catalogage potentiellement récupérable. Des estimations s'appuyant sur une étude réalisée par Ed Jones⁷ indique que plus de 12000 notices bibliographiques pourraient être concernées par an. Les bénéfices réalisées par l'utilisation du fichier AAAF

justifient les investissements et sont directement dûs au partage d'une même culture "catalographique" et aux nombreux recouvrements en matière d'édition entre le Royaume Uni et les Etats-Unis.

Deuxième étude de cas : le Projet AUTHOR

Le Projet AUTHOR illustre bien comment on réalise des performances en rendant accessibles les données d'autorité, comme l'IFLA le recommande. Puisque Françoise Bourdon va présenter le Projet en détail, je me contenterai d'expliquer comment la British Library voit le Projet AUTHOR. Chaque partenaire aura accès via INTERNET aux fichiers d'autorité des autres partenaires. La British Library ne peut pas incorporer directement les notices d'autorité provenant de ces sources dans son propre fichier d'autorité. Résoudre les difficultés techniques que cela pose coûterait trop cher, et, ce qui est plus important, ni la forme du nom ni la langue de catalogage ne serait acceptable dans les catalogues de la British Library. Ceci par opposition avec la culture catalographique partagée avec la Library of Congress. Le coût de fonctionnement des systèmes et du travail d'édition requis pour intégrer de telles notices l'emporterait dans ce cas sur les bénéfices parce que, comparativement, la British Library acquiert peu de documents dans les langues de l'Europe de l'Ouest. Cependant les informations qui aident par exemple à identifier un auteur français, ou à éclaircir les relations entre des collectivités espagnoles sont plus facilement rassemblées par les agences bibliographiques nationales qui en sont responsables que par la British Library. Actuellement de telles informations ne sont pas facilement accessibles parce que les fichiers d'autorité nationaux sont publiés séparément sur différents supports, ce qui rend coûteux leur acquisition, leur gestion, et leur interrogation. Le Projet AUTHOR offre la perspective d'une interface commune qui rendra la British Library capable de réutiliser le travail fait par les collègues des autres pays européens.

Conclusion

Avec cette communication j'ai essayé de mettre en relation l'échange d'information sur les autorités et une meilleure productivité. Une meilleure information sur les gains de productivité que l'on peut espérer est nécessaire avant d'aller plus loin dans la coopération internationale en matière de données d'autorité. Une telle coopération peut se faire à des coûts différents et un gros investissement ne se traduira pas nécessairement par une bonne recette. Certaines conditions doivent être d'abord remplies avant que de gros bénéfices puissent être réalisés. L'IFLA a fait un travail précieux en recommandant des normes pour les données d'autorité, mais la normalisation des règles de catalogage et des formats n'est qu'un préalable à une intégration qui aura lieu ultérieurement. Quand même, il est peu probable que l'IFLA ait le pouvoir d'uniformiser les langues - bien que nombre d'Anglo-Saxons ne parlant qu'une seule langue puisse le désirer en secret - et la langue est un obstacle terrible à l'échange des autorités. Echanger des données à travers les frontières linguistiques coûtera particulièrement cher. Toutefois une approche moins ambitieuse que celle du Fichier d'autorité anglo-américain suffira dans beaucoup de cas. Pour faciliter le travail des collègues dans les autres bibliothèques, les agences bibliographiques nationales devraient, quand cela est possible, suivre l'exemple de la Library of

Congress et rendre leurs données d'autorité accessibles via INTERNET⁹. L'accès à distance à plusieurs fichiers d'autorité, comme le propose le projet AUTHOR, est la forme de coopération la plus économique et peut constituer une plate-forme pour une intégration future. Les agences bibliographiques nationales qui ont une langue en commun pourraient souhaiter réaliser ensemble un fichier d'autorité commun, dans lequel chaque entité bibliographique aurait une forme préférentielle. A long terme, le numéro international normalisé pour les données d'autorité (International Standard Authority Data Number = ISADN) pourrait fournir à chaque entité un identifiant unique indépendant de la langue, des règles de catalogage et de la forme de la vedette. L'ISADN pourrait lier entre elles les grappes de notices d'autorité constituées par langue, et ainsi faciliter l'échange de notices bibliographiques et rationaliser les recherches dans différents catalogues.

Bibliographie

- 1 *Guidelines for authority and reference entries*, recommended by the Working Group on an International Authority System (London : IFLA International Programme for UBC, 1984)
- 2 *UNIMARC/ Authorities : Universal Format for Authorities*, recommended by the IFLA Steering Group on a UNIMARC Format for Authorities (München : K. G. Saur; 1991) (UBCIM Publications - news series, 2)
Note du traducteur : il existe une traduction française :
UNIMARC/Autorités : format universel pour les autorités / IFLA-UBCIM. - Edition française / établie par la Bibliothèque nationale de France. - Paris : Bibliothèque nationale de France, 1996
Cette traduction contient en annexe un tableau complet des zones du format (p.121 à 186). Cette annexe est propre à l'édition française et a été ajoutée avec l'autorisation de l'IFLA-UBCIM. .
- 3 Bourdon, Françoise *International cooperation in the field of authority data* (München : K. G. Saur; 1993) (UBCIM Publications - news series, 11)
Note du traducteur : une version française est disponible gratuitement. La demander à l'auteur à la Bibliothèque nationale de France (2 rue Vivienne, 75084 Paris cedex 02 FRANCE)
- 4 Tillett, Barbara "21st century authority control : what it is and how do we get there?" In : *The future is now : reconciling change and continuity in authority control*. Proceedings of the OCLC Symposium, ALA Annual Conference, June 23rd, 1995 - Dublin, Ohio : OCLC, 1995, p. 17-21.
- 5 Lu, Suping "A study in the Chinese romanization standard in Libraries" *Cataloguing & Indexing Quarterly*, Vol 21 (1) 1995
- 6 Agar, Sharon *Survey of hit rate using an Anglo-American Authority File*. British Library. Unpublished report, 1994.

- 7 JONES, E. A. *Consistency in choice and form of main entry, 1982 and 1989 : a comparison of Library of Congress monograph cataloging with that of the British Library and of the national libraries of Australia and Canada.* (Ph.D. dissertation, University of Illinois at Urbana-Champaign, 1994).
- 8 Jacobowitz, Neil A. "A comparison of AACR2R and French cataloguing rules", *Cataloguing and classification Quarterly*, Vol. 20 (1) 1995
- 9 URL = <http://lcweb.loc.gov/z3950/mums.html>