



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 155-F
Meeting: 86. Bibliography
Simultaneous Interpretation: Yes

Le futur rôle des bibliographies nationales (électroniques)

Juha Hakala

Bibliothèque universitaire d'Helsinki
Bibliothèque nationale de Finlande

Résumé

De plus en plus les bibliographies nationales sont disponibles sur Internet et ont affaire avec le contenu de l'Internet (du Web). Ce texte analyse l'impact que ces changements ont ou devraient avoir sur le rôle des bibliographies nationales. En ce qui concerne l'accès Internet, le recours au protocole Z39.50 est présenté comme étant essentiel pour faire du catalogage par copie. Le catalogage des ressources du Web est examiné à la fois du point de vue des systèmes traditionnels de bibliothèques, de celui des portails de recherche d'information et de celui des systèmes de gestion des objets numériques.

Introduction

En janvier 2004 la Finlande deviendra le 3^e pays au monde à rendre sa bibliographie nationale disponible uniquement sur Internet. La publication de la bibliographie nationale sur support imprimé a cessé il y a déjà quelques années, et fin 2003 la version CD-ROM des *Fennica* entrera aussi dans l'histoire. Le CD-ROM était d'abord et avant tout nécessaire pour le catalogage par copie ; depuis que la plupart de nos clients nationaux ou étrangers peuvent faire du catalogage par copie en ligne via une connexion Z39.50, il n'y a plus de raison de continuer la publication sur CD-ROM. Beaucoup d'autres pays vont être rapidement dans la même situation ; mise à part le respect de la tradition, il n'y a plus aucune raison de continuer à publier la bibliographie nationale sous quelque forme matérielle que ce soit.

Dans cette présentation je voudrais tout d'abord examiner quels sont les pré-requis à l'interruption de la publication de la bibliographie nationale sous une forme tangible. Puis, en utilisant comme point de départ les expériences et les projets finlandais, j'aborderais les différentes façons selon lesquelles les bibliographies nationales peuvent se développer, tout en

gardant à l'esprit l'investissement croissant fait pour le catalogage des ressources électroniques. Dans ce contexte l'implémentation du portail de recherche d'information national finlandais et le système de gestion des objets numériques des bibliothèques universitaires auront un impact important sur la façon dont la bibliographie nationale est produite et sur les services qu'elle peut fournir.

Accès aux bases de données des bibliographies nationales

Selon l'enquête menée par Unni Knutsen, 23 bibliographies nationales étaient disponibles sur Internet en 2001. De plus, 15 agences bibliographiques nationales projetaient de fournir un accès Internet avant 2006 ; parmi elles, au moins 9 ont déjà atteint ce but avant le congrès IFLA de 2003 (Knutsen). Le nombre de bibliographies nationales sur Internet croît très rapidement, alors même que l'accès Web devient le mode d'accès privilégié.

Bien sûr ceci est un développement très positif, mais il manque encore quelque chose. Aussi longtemps que les bibliographies nationales ne seront disponibles sur le Web que pour la recherche, nous ne pourrions pas être pleinement satisfaits, parce que le catalogage par copie est impossible à moins que le protocole de recherche de l'information Z39.50 ne soit implémenté. Tous les efforts que nous avons fait en matière d'harmonisation du catalogage ne deviendront vraiment utiles que si nous rendons nos données disponibles aux autres d'une manière performante ; ceci exige des accords à la fois techniques et politiques.

Examinons en détails les bibliographies nationales électroniques pour déterminer si le catalogage par copie est possible ou non. Pour cela il faut poser quelques questions supplémentaires aux agences bibliographiques nationales qui fournissent un accès Internet :

1. Z39.50 est-il implémenté ?
2. Si oui, dans quel(s) format(s) et avec quel(s) jeu(x) de caractères les données sont-elles fournies ?
3. L'accès Z39.50 est-il libre ou bien est-il nécessaire de passer un contrat ? S'il y a un contrat, y a-t-il des droits à payer ?
4. Seriez-vous favorable à l'établissement de contrats d'échange mutuel de données avec d'autres agences bibliographiques nationales, selon lesquels aucune des parties n'aurait à payer quelque chose ?
5. Si l'accès Z39.50 est impossible, votre bibliographie nationale est-elle en tout ou en partie disponible via un autre système ou d'autres systèmes pour le catalogage par copie ?
6. Si oui, quelles sont les conditions d'utilisation de cette ou de ces bases de données ?

Les agences bibliographiques nationales qui fournissent un accès Internet peuvent, et même doivent, rendre disponibles leurs notices MARC pour le catalogage par copie, et ce en utilisant Z39.50. De très grands catalogues collectifs le font déjà ; par exemple WorldCat (OCLC) qui contenait fin 2002 près de 50 millions de notices, dont environ 60000 décrivent des ressources publiées en finnois. Mais normalement nos bibliographies nationales ont une couverture plus large que celle des catalogues collectifs internationaux, et la qualité des notices est meilleure. Quiconque cherche vraiment un bon recensement des publications finlandaises ou simplement des notices de très bonne qualité, sera plus satisfait en consultant la bibliographie nationale finlandaise. La situation pourrait changer bien sûr si (et seulement si) les notices des *Fennica* étaient chargées dans WorldCat par exemple. Il semble que peu d'agences bibliographiques nationales aient choisi cette stratégie pour le moment ; il serait

intéressant de savoir si certaines d'entre elles ont des projets de coopération plus étroite avec OCLC par exemple.

Z39.50 et les bibliographies nationales

Z39.50 est aujourd'hui une norme qui caractérise tout bon système intégré de bibliothèque et tout bon portail de recherche d'information, tel que MetaLib. Le problème est que, malgré l'existence de principes directeurs approuvés au niveau international pour l'implémentation de Z39.50 - le profil de Bath (Bath) - les systèmes intégrés de bibliothèque ne se "parlent" pas encore facilement les uns les autres. Il y a beaucoup de raisons à cela, y compris les suivantes :

- au niveau international toutes les bibliographies nationales n'utilisent pas MARC21 (qui est le format requis pour le profil de Bath), et parmi celles qui n'utilisent pas MARC21, certaines ne peuvent pas convertir leurs données en MARC21 à la volée pour un échange de données basé sur Z39.50. Puisqu'une conversion manuelle de format n'est pas faisable pour un catalogueur qui souhaite récupérer les notices, c'est un sérieux problème. La meilleure façon de le régler est d'utiliser un bon convertisseur de formats MARC, tel que USEMARCON (USEMARCON) comme élément du processus de catalogage par copie. Alors, même si la notice récupérée arrive dans un "mauvais" format MARC, la conversion vers le format MARC local peut être facilement faite. Tout ce dont vous avez besoin pour réaliser cela c'est un jeu de tableaux de conversion qu'un expert en format, connaissant bien à la fois le format source et le format cible, pourra construire en quelques semaines avec USEMARCON.
- nos formats MARC utilisent une grande variété de jeux de caractères, et il arrive que des notices en MARC21 par exemple, soient encodées de différentes façons. UNICODE changera cela éventuellement - peut-être - mais beaucoup d'eau passera dans la Spree avant que tous les systèmes de bibliothèque ne prennent complètement en compte UNICODE. Z39.50 propose une solution intelligente à ce problème : la négociation du jeu de caractères. Mais seulement quelques systèmes de bibliothèque peuvent prendre en compte cette caractéristique dans le service Initialisation de Z39.50. L'utilisation de USEMARCON ou d'une autre application capable de convertir un jeu de caractères d'une manière adéquate est la solution la plus rapide à ce problème.
- différents fournisseurs Z39.50 produisent différentes sémantiques (c'est-à-dire des critères de recherche différents). Le profil de Bath donne des principes directeurs explicites sur comment construire de façon normalisée, dans un contexte Z39.50, des requêtes et des combinaisons de critères de recherche. Il n'y a pas de raison pour que les bibliothèques nationales s'écartent des recommandations du profil ; la seule exception est que, étant donné la richesse des éléments de données dans la bibliographie nationale, il est possible, sinon recommandé, de fournir des points d'accès supplémentaires. Mais même alors, la norme Z39.50, et plus particulièrement son ensemble d'attributs de recherche Bib-1 devrait constituer une définition a minima de ce qui doit être offert. Heureusement on peut aller au-delà de ce minimum, puisqu'il est très facile de définir de nouveaux termes dans Bib-1.
- différents fournisseurs et clients Z39.50 implémentent différents ensembles de services et différents paramètres pour ces services, tous définis dans la norme. Encore

une fois, le respect du profil de Bath est le meilleur moyen d'éviter les problèmes dans ce domaine. Heureusement le catalogage par copie ne requiert rien d'exotique ; les services de base Init, Search et Present sont suffisants.

Pour s'appuyer sur un exemple pratique, examinons la situation en Finlande. Notre bibliographie nationale est en MARC21-Fin, version finlandaise étendue et modifiée de MARC21. Les bibliothèques universitaires et de recherche utilisent le même format, mais les bibliothèques publiques utilisent encore le vieux format FINMARC et le jeu de caractères ISO 6937/2. Cependant les bibliothèques publiques n'ont pas de difficulté à utiliser les *Fennica* pour le catalogage par copie.

Les plus grands fournisseurs de systèmes de bibliothèque finlandais ont intégré USEMARCON dans leurs modules de catalogage (et ont amélioré l'application à cette occasion). Quand une bibliothèque publique utilise un de ces systèmes de récupération de notices en format MARC21-Fin à partir de la bibliographie nationale via Z39.50, USEMARCON convertit ces notices à la volée en FINMARC et ISO 6937/2. Une des conséquences prévisibles de cette procédure a été que l'accès en ligne est devenu très populaire, tandis que les abonnements aux *Fennica* sur CD-ROM diminuaient. Donc la décision de mettre fin au CD-ROM *Fennica* a été facile à prendre ; dans notre cas, seule la version en ligne de la bibliographie nationale est utile. Un autre facteur favorisant est le développement phénoménal de l'informatique ; il y a 10 ans il aurait été impossible d'acheter un serveur suffisamment puissant pour permettre le catalogage par copie dans toutes les bibliothèques publiques finlandaises ; aujourd'hui ce n'est plus du tout un problème.

En 2004 nous offrirons un accès Z39.50 aux usagers étrangers, avec une fourniture de données en MARC21. USEMARCON a aussi été intégré dans le serveur Z39.50 de Voyager, et ceci nous permet, au moins en théorie, de fournir des notices bibliographiques dans n'importe quel format MARC avec le jeu de caractères correspondant. Dans la pratique, seulement une poignée de formats peuvent être maintenus par une agence bibliographique nationale donnée, puisque la maintenance des tableaux de conversion nécessite une parfaite maîtrise à la fois du format source et du format cible. En utilisant MARC21 comme commun dénominateur, les agences bibliographiques nationales pourraient fournir un grand éventail de conversions, permettant ainsi à toute bibliothèque nationale de récupérer des notices MARC directement de n'importe quelle bibliographie nationale via Z39.50.

L'utilisation de Z39.50 complétée par un logiciel de conversion de formats et de jeux de caractères tel que USEMARCON, permet une recherche de l'information et une récupération de notices tout à fait efficaces à partir des bibliographies nationales. Puisque la technologie est maintenant bien installée, il est urgent de réviser les recommandations sur les services bibliographiques nationaux y compris les principes directeurs du Z39.50. Il est curieux que cette norme, ou sa version ISO, ISO 23950, ne soit pas même mentionnée dans les recommandations de la Conférence internationale sur les services bibliographiques nationaux de 2002 (i.e. 1998 : note du traducteur).

Une fois que l'infrastructure technique est en place, il faut créer les contrats nécessaires à l'échange des données entre agences bibliographiques nationales. Une coopération internationale telle que le Catalogue collectif virtuel scandinave (Scandinavian Virtual Union Catalogue <<http://www.lib.helsinki.fi/svuc/>>) liant les catalogues collectifs nationaux nordiques en un unique service virtuel, ou l'Association ONE <<http://www.oneassociation.org>> ont pour but de rendre accessibles les données

bibliographiques au niveau national et international, en prenant souvent en compte les notices des bibliographies nationales incluses dans un catalogue collectif.

Puisque les agences bibliographiques nationales échangent assez peu leurs notices bibliographiques de façon bilatérale, le partage de ces données devrait être fait dans un contexte plus large. L'Association ONE pourrait, je crois, servir de modèle pour des initiatives de partage de notices. D'un autre côté, le partage des données d'autorité entre agences bibliographiques nationales est vraiment pertinent, et nous devons faire de notre mieux pour construire ce qu'on appelle des Fichiers d'autorité internationaux virtuels. Mais alors même que la version 2.0 du profil de Bath spécifie vraiment comment les données d'autorité peuvent être recherchées via Z39.50, les systèmes de bibliothèque ne disposent pas encore d'implémentations Z39.50 présentant cette fonctionnalité - pour ne pas mentionner d'autres avancées techniques et politiques absentes pour le moment.

Comme le montre l'exemple de l'échange des notices d'autorité, les bibliothèques nationales semblent avoir des spécifications liées à un Z39.50 spécial et bien d'autres spécifications que les autres bibliothèques ont pour règle de ne pas avoir, et nous voudrions tranquillement former notre propre groupe d'utilisateurs d'un fournisseur spécifique dans le but d'imposer le développement de caractéristiques qui nous sont indispensables. Un tel groupe existe déjà : le National Libraries' Endeavor Advisory Board, le NLEAB <voir <http://www.infoday.com/IT/nov01/news18.htm>>.

Que faire du protocole ZING ?

Beaucoup d'entre vous ont entendu parler du protocole ZING (ZING). Cet acronyme signifie Z39.50 International Next Generation, la nouvelle génération internationale du Z39.50. Techniquement il est tout à fait différent du Z39.50, encore que pour l'utilisateur les choses semblent se passer à peu près de la même manière. L'agence de maintenance du ZING est la Bibliothèque du Congrès, qui assure la coexistence pacifique de ces normes.

Pour dire les choses simplement, le protocole ZING appliquera Z39.50 au Web. En utilisant par exemple un service ZING SRU, il est possible d'envoyer des requêtes encodées dans les URLs, de la façon suivante :

<http://www.loc.gov/sru.cgi?query=dc.title=cat&maximumRecords=10&recordSchema=DC&sortKeys=title,dc&startRecord=1>

Bien que les données paraissent épouvantables, ça fonctionne ! Par exemple Google utilise une solution technique similaire, et de même dans un OpenURL les données bibliographiques relatives à une ressource envoyées à l'application de résolution des OpenURL sont entassées dans un URL. Le serveur ZING renvoie les notices non pas en format d'échange ISO 2709 mais encodées en XML sur la base d'un schéma Dublin Core.

Puisque la plupart des fournisseurs de systèmes de bibliothèque ont déjà implémenté Z39.50 pour lequel le nombre de services définis est beaucoup plus grand que dans la présente version du protocole ZING, ce ne serait pas une bonne idée de remplacer Z39.50 par ZING dans les systèmes de bibliothèque. Mais il peut s'avérer utile d'implémenter à la fois Z39.50 et ZING pour rendre les systèmes de bibliothèques plus accessibles en-dehors du secteur des bibliothèques. Par exemple, si une bibliographie nationale était disponible via ZING, il serait

plus facile pour les éditeurs et les autres établissements de mémoire tels que les musées ou les centres d'archives, de connecter leur système à la bibliographie nationale.

Selon moi, les agences bibliographiques nationales devraient encourager l'utilisation des données des bibliographies nationales en rendant leurs notices disponibles en syntaxe XML via le protocole ZING et si possible aussi via le protocole OAI (Open Archives Initiative <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>). Dans quelques années nos systèmes de bibliothèques posséderont cette fonctionnalité ; en attendant nous devons nous préparer à utiliser ces technologies.

Les futurs rôles des bibliographies nationales

Selon l'enquête de Unni Knutsen présentée tout à l'heure, les bibliographies nationales peuvent ou non être fusionnées dans l'OPAC de la bibliothèque nationale ou d'autres systèmes bibliographiques, mais il y a une chose que nous avons jusqu'à présent considérée comme admise : la bibliographie nationale est élaborée grâce à un système intégré de bibliothèque mais d'un point technique elle constitue vraiment *une* base de données.

A partir de là, le système intégré de bibliothèque tel qu'il est habituellement connu est sérieusement handicapé. Des fournisseurs de systèmes de bibliothèque et d'autres entreprises sont en train de développer des portails de recherche d'information - MetaLib chez Ex-Libris étant probablement l'exemple le plus connu de ces applications - ainsi que des systèmes de gestion d'objets numériques ou DOMS (Digital Object Management Systems) tel que ENCompass chez Endeavor Information Systems.

On peut considérer les portails de recherche d'information comme une grande extension d'un OPAC traditionnel. Un portail permet généralement de rechercher simultanément dans plusieurs bases de données cibles, en utilisant différents protocoles de requêtes, différentes syntaxes de notices et différents jeux de caractères. Il permet aux clients d'archiver les requêtes et de demander au système de les relancer périodiquement. Dernièrement, le fait que les portails utilisent des liens dynamiques avec des OpenURL est aussi devenu une "norme". Bien que ces services puissent être intégrés dans un système traditionnel de bibliothèque, les fournisseurs ont souvent construit des applications séparées pour eux. Il y a à cela tout un tas de raisons techniques et commerciales ; une bibliothèque utilisant un système intégré venant d'un fournisseur A peut acheter un portail chez un fournisseur B sans avoir à changer de système.

Le même contexte technique et commercial s'applique aux DOMS : un système de bibliothèque pourrait incorporer la fonctionnalité nécessaire, mais puisque un bon DOMS doit pouvoir gérer une grande variété de formats de documents et de métadonnées et être capable d'effectuer une fourniture contrôlée de contenus protégés, de nombreux développements sont nécessaires pour qu'un système traditionnel de bibliothèque puisse acquérir ces fonctionnalités.

Quand les bibliothèques installent des systèmes de portails ou de DOMS, un système intégré unique peut être remplacé par deux ou trois applications, interconnectées grâce à des interfaces standard. Dans le réseau finlandais Linnea, à partir de 2004, nous utiliserons Voyager comme système de bibliothèque, MetaLib comme portail de recherche d'information et ENCompass comme DOMS.

Du point de vue du catalogage, l'émergence de ce nouvel environnement technique est intéressant pour deux raisons. D'abord l'information sur les liens qui est vitale pour la description des ressources (électroniques) pourrait ne plus être du tout donnée dans le système traditionnel de bibliothèque, mais dans le portail. D'autre part, l'utilisation efficace d'un portail et d'un DOMS va nécessiter de nouveaux types de métadonnées pour lesquelles nous n'avons encore ni règles de catalogage ni formats. Ces évolutions vont avoir un impact à la fois sur les OPAC et sur les bibliographies nationales.

Permettez moi de décrire plus à fond le premier problème. Une solution couramment utilisée pour fournir un lien entre un système basé sur MARC21 et une ressource Web consiste à indiquer l'URL de la ressource Web dans le champ 856 de la notice MARC. Puis, chaque fois que l'URL change nous devons mettre à jour la notice dans chaque base de données où elle a été récupérée. Malheureusement, en Finlande du moins, des bases de données bibliographiques contiennent assurément quelques URL qui ne fonctionnent plus, ou pire encore, qui établissent un lien entre la notice MARC et une ressource différente de celle qui a été cataloguée au départ.

La solution reposant sur le champ 856 est aussi incapable de régler la question de la copie adéquate ; c'est-à-dire qu'on peut multiplier les copies [les exemplaires] d'une même ressource à différentes adresses de sorte que tel utilisateur soit autorisé à utiliser l'une de ces copies, tel autre une autre copie, et un troisième encore une autre copie. Parce qu'il est statique le système de lien basé sur une notice MARC ne peut tout simplement pas prendre en compte ces situations qui vont devenir de plus en plus courantes avec les systèmes de licences d'utilisation.

Le cas de la copie adéquate peut être réglé, et la maintenance des URL simplifiée, en implémentant les OpenURL. Avec cette approche, l'information sur les liens et les licences d'utilisation est stockée dans la base de données de résolution des OpenURL ; en Finlande il s'agit du système SFX incorporé dans le portail Metalib. Une fois que cette opération est réalisée, on peut retirer de la bibliographie nationale les données sur les liens parce que ce serait contre-productif de les y laisser. Cela n'aurait pas de sens de maintenir à la fois des liens statiques dans la bibliographie nationale et des liens dynamiques dans la base de résolution des OpenURL.

Notre conclusion est donc que la bibliographie nationale soit divisée entre un système traditionnel de bibliothèque et une application de résolution des OpenURL telle que SFX. Des métadonnées relatives par exemple à la préservation à long terme peuvent, elles, être placées dans un système de gestion des objets numériques. Une question logique se pose alors, quelles sont les métadonnées utilisées par l'application de résolution des OpenURL que l'on peut considérer comme appartenant à la bibliographie nationale (1) ?

Beaucoup de catalogueurs probablement vont considérer que, si des éléments de données existants sont déplacés ou copiés du système de bibliothèque dans le système de résolution des OpenURL, alors cette information continue d'appartenir à la bibliographie nationale. Le contenu du champ 856 en est un bon exemple. Seuls les éléments de données qui ne sont pas encore utilisés restent problématiques. Pour dire les choses autrement, la bibliographie nationale devrait-elle par exemple contenir des informations sur les licences d'utilisation des ressources électroniques ? Ma réponse à cette question est un "oui mais" ; si nos systèmes se développent de telle manière qu'ils soient capables de stocker et d'utiliser les informations sur les licences d'utilisation, nous ne devrions utiliser ces données que lorsqu'elles sont assez

stables. Par exemple les modalités et les conditions d'utilisation applicables aux ressources déposées devraient assurément être indiquées dans la bibliographie nationale.

Ainsi, dans un futur proche, nos bibliographies nationales pourront être constituées d'un système traditionnel de bibliothèque et d'une application de résolution des OpenURL. Du point de vue de l'utilisateur final, ces deux applications constituent un service unique, fournissant des liens dynamiques aux contenus électroniques. Ces contenus peuvent être stockés dans un système de gestion d'objets numériques qui constituera le troisième volet de nos futurs services. Nous ne mesurons pas encore complètement toutes les implications de ce nouvel environnement applicatif ; par exemple, nous ne savons pas encore fournir des notices bibliographiques en même temps que les informations sur les liens si les données ont été réparties sur deux applications. Mais ces questions en suspens seront à discuter et éventuellement à résoudre dans un cadre tel que celui de l'IFLA.

Note

(1) Bien sûr, les applications de résolution des OpenURL peuvent être intégrées aux systèmes intégrés traditionnels de bibliothèque, et les opérations de gestion peuvent devenir des routines dans certains systèmes d'information. A ce moment là on peut encore avoir un système intégré unique, peut-être même un seul format ou des formats appartenant à la même famille, capables de prendre en compte toutes les métadonnées nécessaires. Mais que cette intégration existe ou non, notre conception de la bibliographie nationale doit être élargie.

Références bibliographiques

Bath profile home. The National Library of Canada [en ligne]. <http://www.nlc-bnc.ca/bath/> (2003-07-10)

Knutsen, Unni : Bibliographies nationales électroniques : point sur la situation actuelle [en ligne]
http://www.ifla.org/IV/ifla69/papers/109f_trans-Knutsen.pdf (2003-07-10)

OCLC WorlCat home [en ligne]. OCLC Online Computer Library Center
<http://www.oclc.org/worldcat/> (2003-07-10)

USEMARCON home [en ligne] The British Library.
<http://www.bl.uk/services/bibliographic/usemarcon.html> (2003-07-10)

Z39.50 Maintenance agency home page. Library of Congress [en ligne]
<http://lcweb.loc.gov/z3950/agency/> (2003-07-10)

ZING home page. Library of Congress [en ligne]
<http://lcweb.loc.gov/z3950/agency/zing/zing-home.html> (2003-07-10)