

Cataloguer les cédéroms et autres documents électroniques

par Dominique Lahary*

Les documents électroniques et notamment des cédéroms sont présents dans un nombre croissant de bibliothèques. Ils posent certes des problèmes technique (mise à disposition sur place, responsabilité vis-à-vis des emprunteurs d'un produit encore peu stabilisé) et juridiques (a-t-on le droit de prêter?). Mais ils posent aussi des problèmes de traitement bibliographique, et l'on entend de plus en plus : « *Comment faites-vous pour les cataloguer ?* » Le but du présent article est de donner quelques indications simples et pratiques.

Des normes...

On sait que les ISBD sont publiés par type de document, cette notion vague correspondant tantôt à la nature de l'information (documents sonores), au mode de publication (les publications en série) ou au support matériel. C'est ainsi qu'en 1988 a été publié l'ISBD (CF) relatif aux fichiers électroniques (*computer files*), remplacé en 1997 par l'ISBD(ER), la notion de ressource électronique (*electronic resources*) permettant une application plus large, notamment aux ressources en ligne. La norme AFNOR correspondante, qui porte la référence Z 44-082, devrait être publiée en 1998.

Qu'ont donc les documents électroniques de si particulier du point de vue du catalogage ? Peu de choses en vérité. La première caractéristique est une banalité : la source des informations se trouve sur le document lui-même. Tout bibliothécaire cataloguant un imprimé à partir de sa page de titre sait cela. Mais cela pose un problème pratique : ouvrir un cédérom, c'est plus long qu'ouvrir un livre. Et certaines mentions comme celles relatives à la responsabilité figurent parfois sur des écrans furtifs ou déroulants.

Après le titre propre, les normes autorisent facultativement pour tous les non livres de faire figurer entre crochets carrés l'indication générale du type de document : ce sera donc *Ressource électronique*. Dans un catalogue informatisé, cette mention alourdit l'affichage et n'est souvent là que parce que le système ne sait pas indiquer autrement le type de document, quand d'autres déjà gèrent cela par icône.

Les mentions de responsabilités peuvent poser problème pour de nombreux cédéroms du fait de leur multiplicité. On retrouve dans les œuvres multimédias le phénomène de l'excès d'auteurs ou de collaborateurs déjà constaté dans le cinéma et la vidéo. On pourra s'en tenir à l'essentiel.

La notion de mention d'édition correspond souvent, dans le cas des documents informatiques, à celle de version, par exemple

version d'un logiciel.

La zone 3 de l'ISBD, dite *spécifique à certains types de document*, existe pour les documents électroniques. Elle est utilisée pour indiquer quel est le contenu de document. Il s'agit de *données* (on peut préciser : *textuelles, représentationnelles, iconographiques, vidéographiques, sonores...*) et/ou de *logiciels* (on peut préciser : *logiciel utilitaire, d'application, système, pilote*). On peut simplement noter *Données et logiciels, Multimédia interactif* (dans le cas des CD-ROM multimédias), enfin *Service en ligne*. On peut enfin donner entre parenthèses des indications sur la taille des données : *Données informatiques (3 fichiers : 800 enregistrements, 3 150 octets) et programmes (2 fichiers : environ 650 instructions chaque)*.

La collation permet d'indiquer la nature et le nombre d'unités matérielles, ainsi que les éventuels matériels d'accompagnement. Cette zone n'existe pas pour les services et données en ligne.

Quant aux notes spécifiques à ce type de document, elles sont au nombre de trois et présentent un caractère obligatoire. La première poste sur la source du titre. La seconde concerne les supports physiques comme les cédéroms ou les disquettes. Elle commence par la mention « *Configuration requise* : » et comprend, dans l'ordre, le type de machine, la mémoire vive nécessaire, le nom et les versions des systèmes d'exploitation et interfaces gra-

* Conservateur à la BDP du Val d'Oise.

phiques, éventuellement des détails sur le logiciel et le langage de programmation, enfin les périphériques requis.

Cette note est répétable si plusieurs configurations sont possibles (par exemple PC et Macintosh). Enfin les ressources en ligne requièrent une note sur l'accès, comprenant toutes les informations nécessaires comme le protocole, le chemin d'accès, le nom de fichier, etc.

...aux formats

Nous nous bornerons à l'exemple d'UNIMARC, format le plus connu en France. Ce format est par principe fait pour décrire tout type de ressource, il ne pose donc pas de difficulté particulière et la dernière édition disponible en français offre toutes les données nécessaires, à l'exception de la zone 856.

Le format, en son état actuel, souffre notamment de deux manques : la liste des codes de fonction ne couvre pas la variété des contributions de la création informatique et multimédia, et la numérotation commerciale, non normalisée (et parfois redoublée par un ISBD) ne fait l'objet d'aucune zone. Certains utilisent, à défaut, la zone 071 en principe réservée (mais on se demande bien pourquoi) à la numérotation non normalisée des documents sonores.

Une banalité : la source des informations se trouve sur le document lui-même. Mais cela pose un problème pratique : ouvrir un cédérom est plus long qu'ouvrir un livre...

Les seules zones spécifiques aux ressources électroniques sont les suivantes :

135 : Données codées

Une seule sous-zone et une seule position : il s'agit d'une typologie sommaire (7 valeurs possibles représentées par une lettre).

230 : Zone spécifique à certains types de document

Cette zone correspond à la zone 3 de l'ISBD.

336 : Note sur les fichiers informatique

L'utilité de cette zone n'apparaît guère car elle est largement redondante avec la zone 230.

337 : Note sur la configuration requise ou sur l'accès

Cette zone est naturellement essentielle. Son contenu a été détaillé ci-dessus.

626 : Accès par les données techniques

Cette zone sert à procurer un point d'accès «pour l'accès en ligne ou pour sélectionner et trier les notices en vue d'un index imprimé». Sont successivement indiqués en Sa, Sb et Sc : ma marque et le modèle de machine, le langage de programmation, le système d'exploitation. A ne pas confondre avec la note sur la configuration, faite pour être affichée.

Cinq zones non spécifiques aux documents électroniques peuvent être citées :

Le guide (ou en-tête, ou label, ou leader)

La position 6, qui fait l'objet d'une saisie guidée dans la plupart des logiciels de bibliothèque, doit comporter la valeur *l* (support informatique) et non *m* (document multimédia). La seconde édition française du manuel UNIMARC utilise encore malheureusement ce terme dans le

sens aujourd'hui compris des seuls bibliothécaires de *multi-support* (en anglais *kit*).

200 : Zone du titre et de la mention de responsabilité

C'est la sous-zone Sb qui peut contenir, si l'on y tiens, la mention *Ressource électronique*.

304 : Note sur la zone du titre et de la mention de responsabilité

Elle contient la note obligatoire sur la source du titre.

315 : Note sur les informations propres au type de document

On peut y mentionner toute information utile en complément à la zone 230.

856 : Lien à un document électronique

Cette zone nouvellement introduite dans UNIMARC¹ permet de lier une notice à une ou plusieurs ressources électroniques accessibles en ligne. Si elle permet de lier ainsi la notice d'une monographie imprimée à l'image de la page de couverture, au sommaire numérisé, etc., elle est naturellement toute indiquée pour une notice décrivant un document accessible en ligne, et réalise (si le logiciel de gestion de catalogue le permet) le lien entre la description et la ressource elle-même. Seuls les OPAC graphiques, en particulier ceux basés sur le World Wide Web (Intranet et/ou Internet), permettent d'activer de tels liens.

Récupérer plutôt que cataloguer ?

Cataloguer un CD-ROM peut-être long, surtout si l'on respecte la règle qui veut que les informations doivent être tirées du document lui-même et non des conteurs et imprimés d'accompagnement. L'intérêt de la récupération de l'information bibliographique, qui vaut pour tout type de document, est décuplé en ce qui concerne les documents électroniques.

Malheureusement, la situation actuelle n'est guère satisfaisante : il n'y a pas assez de notices récupérables en format d'échange, et elles sont inutilement dispersées. ce qui explique que les bibliothèques créent la plupart elles-mêmes leurs notices de CD-ROM, selon une courte enquête que nous avons menés notamment auprès de membres de l'ADDNB².

1. Voir La zone 856 : «Le lien au document primaire», par Dominique Lahary. *Bulletin d'informations de l'ABF* n°174 et «Qu'est-ce que le Comité permanent UNIMARC ?», par Elisabeth Freyre, *Bulletin d'informations* n°175.

2. ADDNB : Association pour la diffusion des documents numériques en bibliothèques. Bibliothèque municipale 33 rue du Gouverneur-général-Eboué 92 130 Issy-les-Moulineaux.

La BnF

La Bibliothèque nationale de France diffuse des notices de documents électroniques sur deux produits différents : les cédéroms *BnF-Livres, publications en séries et documents électroniques* et *BnF-Documents sonores, audiovisuels et multimédias*.

Les critères de répartitions sont fort difficiles à comprendre pour l'utilisateur³. La solution préférée de ceux-ci serait sans nul doute le rassemblement de toutes ces notices sur un seul produit, et sur le plus utilisé : *BnF-Livres*, d'autant que le problème des notices à niveau ne se pose guère pour ce type de document.

On relève en outre des différences de traitement entre les notices issues des deux bases BN-OPALE ET BN-OPALINE après conversion en UNIMARC. Dans *BnF-Livres*, le numéro commercial n'est pas fourni, ce qui est orthodoxe mais dommageable. Il l'est dans *BnF-DSAM*, dans la zone 071, mais la configuration requise est fournie en 300 (note générale) au lieu de 337 (note sur les détails techniques), à la différence de l'autre produit.

Enfin, l'examen des dernières éditions disponibles de ces produits révèle un taux de recouvrement insuffisant : il y a 1 354 notices de documents électroniques dans *BnF-Livres*, et seulement 144 dans *BnF-DSAM*, ce qui ne permet guère de couvrir

3. Le CD-ROM *BnF-Livres* comprend plutôt les produits multimédias, mais le premier comporte un certain nombre de documents décrits dans la notice comme multimédias. On trouvera dans l'un ou l'autre les encyclopédies, et tandis que les *100 fiches cuisines de Elle* sont du côté des livres, *La cuisine de bien-être de Valérie-Anne* côtoie la musique.

Le DOI, l'ISBN de l'Internet?

par Dominique Lahary

Le DOI (*Digital Object Identifier*) est un système d'identification des documents accessibles sur Internet. Né d'une initiative de deux organismes américains, l'AAP (American Association of Publishers) et le CNRI (Corporation for National Research Initiatives) et d'un organisme international, le STM (International Association of Scientific, Technical and Medical Publishers), il a reçu le soutien de l'Union internationale des éditeurs.

Une DOI-Foundation¹ a été créée pour le promouvoir. On trouvera sur son site Web une abondante information ainsi que des liens vers les sites d'éditeurs expérimentant le DOI².

Le DOI possède une syntaxe simple et peu rigide : il est composé de deux éléments séparés par une barre oblique. Le premier ou préfixe comprend l'identi-

1. L'éditeur français Jean-Manuel Bourgois est membre de son conseil exécutif.

les besoins de traitement des CD-ROM multimédias pour grand public.

ELECTRE

Le taux de recouvrement est d'une toute autre ampleur dans la base ELECTRE-Multimédia, reproduite sur le CD-ROM bibliographique moyennant un supplément de prix : plus de 8 000 CD multimédias (CD-ROM, CDI et autres variantes) sont référencés. Malheureusement, ces notices ne sont actuellement pas déchargeables. Certaines bibliothèques s'en servent toutefois pour recopier la description bibliographique.

Références bibliographiques

- *ISBD(CF) : description internationale normalisée des documents électroniques = International standard bibliographic description for computer files* / IFLA, UBCIM programme ; trad. par Florence Liégeois sous la dir. de Françoise Leresche. — Paris : Bibliothèque nationale de France, 1996.
- *ISBD(ER) : International standard bibliographic description for electronic resources* / IFLA, UBCIM programme. — München ; London ; Paris : Saur, 1997.
- Le projet français AFNOR Z 44-082 *Catalogage des ressources électroniques*. A paraître en 1998.
- *Manuel UNIMARC : format bibliographique* / IFLA, UBCIM programme. — Version française, 2^e éd. — München ; London ; Paris : Saur, 1996.

Ressources disponibles sur le World Wide Web

- Guidelines for the Use of Field 856: <http://lcweb.loc.gov/marc/856guide.html>
- UNIMARC concise format : <http://ifla.inist.fr/VI/3/p1996->

fiant de l'agence DOI (deux chiffres) suivi de l'identifiant de l'éditeur (?? chiffres). Le second ou suffixe est l'identifiant du document attribué par l'éditeur, sans structure ni longueur imposées.

Qu'est-ce qu'un objet numérique? La définition en est lâche, et le DOI est susceptible de concerner un large éventail d'unités : il peut s'agir aussi bien d'un titre de périodique que d'articles, d'une encyclopédie en ligne que de chacune des entrées, ou encore d'une image.

Mais plus qu'un système d'identification,

2. <http://www.doi.org>. Visiter notamment la Gallery et, au menu « About the DOI », sélectionner « Run through a system demonstration » pour suivre une démonstration. Ce site est évidemment en anglais. On pourra lire en français : « DOI : La norme électronique pour le commerce en ligne de demain » par Daniel García, in *Livres-Hebdo* n°268, 7 novembre 1997.

le DOI est un système de routage, et en définitive de centralisation et de contrôle des accès. Quand l'utilisateur final demande l'accès à un « objet » muni d'un DOI³, sa demande aboutit dans une agence DOI. C'est elle qui connaît la localisation physique du document et transmet la demande de connexion au serveur qui l'héberge. Il y aura un certain nombre d'agences DOI servant ainsi à la fois de gestionnaires des localisations et d'intermédiaires pour l'accès. Dans la phase expérimentale actuelle, une seule agence a été mise en place.

Le DOI permet ainsi d'échapper au caractère instable des adresses du World Wide Web en instituant une identification indépendante de la localisation physique et en

3. Une icône DOI a été créée. Il suffit de cliquer dessus pour accéder au document... ou à la page qui en décrit les conditions d'accès.

confiant à un site intermédiaire le soin du routage. Ce souci était déjà celui qui avait amené l'IETF (Internet Engineering Task Force) à définir la notion d'URN (*United Resource Name*) par opposition à l'URL (*United Resource Locator*)⁴. Les promoteurs du DOI le présentent comme un URN.

Si le DOI n'est pas en lui-même un système de tarification, le contrôle des accès qu'il implique est susceptible de faciliter la mise en place de dispositifs de paiement par abonnement où à l'acte, et plus largement d'interdiction d'accès aux utilisateurs non autorisés. Bien que le site du consortium mette aussi en valeur des utilisations non lucratives du DOI, il pourrait bien s'agir là d'une des principales raisons de son éventuel succès auprès des éditeurs.

4. <http://ietf.org/html.charters/urn-charter.html>

1/concise.htm (en anglais)

• Bordereau de catalogage des CD-ROM en UNIMARC / Bibliothèque départementale du Val d'Oise : <http://www.cg95.fr/biblio/bdvo/catcd r2.htm>

• On trouvera des notices en USMARC de documents électroniques accessibles en ligne dans le catalogue du projet Intecat d'OCLC : <http://orc.rsch.oclc.org:6990/>

(Faire une recherche, afficher une notice, puis cliquer sur le bouton *MARC display*).

(Remerciements à : Françoise Boissière, bibliothèque de l'ENS Cachan ; Edwige David, bibliothèque municipale de Saint-Etienne ; Patrick Gozkowski, bibliothèque départementale du Val-d'Oise et Antoine Provansal, BnF.)

• En bref...

Z39.50 et ses profils d'implémentation européens

par T. Vebr

Le 20 octobre 1997, l'EFILA (European Forum for Implementors of Library Automation¹) a consacré à Bruxelles une séance aux profils européens d'implémentation de la Z39.50. Plusieurs projets nationaux de profils d'implémentation ont été présentés :

* Royaume-Uni : projet MODELS (MOving

to Distributed Environment Library Services) ;

* France : projet du CCF (Catalogue collection de France) ;

* Finlande : profil finnois ;

* Italie : projet du SBN (Servizio Bibliotecario Nazionale)².

Ces profils d'implémentation, qui portent

2. Le catalogue collectif italien avec passerelle Web/Z39.50 est interrogeable à l'adresse suivante : <http://opac.sbn.it>

1. <http://www.dbc.dk/efila>

notamment sur le nombre et la définition des clés de recherche ou sur la façon dont la base cible a répondu à la requête (mot n'importe où dans le champ, formule normalisée, etc.) permettent d'éviter les réponses nulles.

Il n'existe pas de profil européen commun à tous ces projets de bibliothèques nationales, bien qu'ils s'inspirent tous du profil du CENL (Consortium of European National Libraries) et des profils américains.

Le traitement des données locales (holdings) reste un problème pour tous, que ce soit en ce qui concerne les données d'inventaire des monographies, les inventaires globaux ou détaillés des périodiques. Il n'existe pas pour l'instant de solution théorique : la norme Z39.50 ne gère pas cet aspect dans son état actuel (version 3).

Les formats dans lesquelles les notices sont envoyées par le serveur (UNIMARC, USMARC, UKMARC, FINMARC...) demeurent un grand point de dissension. L'implémentation systématique d'un format de sortie autre que le format MARC et lisible par les humains (SUTRS) pourrait être une alternative à ces divergences de format MARC.

Les divergences entre profils proviennent aussi de ce que certains pays désirent intégrer dans leurs catalogues les données des bibliothèques mais aussi celle des musées et des archives.

L'implémentation de la fonction SCAN de feuilletage de l'index (*Browse the index*) s'avère nécessaire, surtout pour les recherches par auteurs ou vedettes matière : en effet il existe de grandes différences selon les bibliothèques entre la structure des formes saisies ou les classifications ou thesauri utilisés.

