

Table ronde

# Les nouveaux domaines de la normalisation

---

Compte rendu des interventions de :

Yves Moret, *chargé de mission pour le CCF (BNF-DINT)*,

Jean-Marc Czaplinski, *BNF-DINT*,

Christine Deschamps, *service commun de documentation de l'université de Paris 5-René Descartes*,

Dominique Lahary, *BDP du Val-d'Oise*,

François Role, *DISTNB*.

Yves Moret, chargé de conduire le projet de Catalogue collectif national à la BNF, introduit le débat en rappelant que les bibliothécaires ont depuis longtemps l'habitude de se débattre avec nombre de normes, mais aussi de standards, comme les formats bibliographiques. Mais voici qu'une nouvelle génération agite les milieux professionnels – ou tout au moins, une partie d'entre eux : Z39.50, SGML...

## **Z39.50, ou comment une norme américaine accède à l'universalité**

Christine Deschamps retrace l'historique de la norme Z39.50. Au commencement était l'OSI (*open system interconnection* = interconnexion des systèmes ouverts), modèle d'interconnexion définissant sept couches (de la connexion physique à l'application) défini par l'ISO (*International Standard Organisation*). La mise au point du modèle technique a permis à

l'ISO de définir deux séries de normes, relatives à des services : ISO 10160 et 10161 pour l'ILL (*Inter Library Loan* = prêt entre bibliothèques), ISO 10162 et 10163 pour *Search & Retrieve* (chercher et trouver). Cela signifie qu'on abordait parallèlement le prêt entre bibliothèques et la recherche sur les catalogues. Dans la tradition des normes internationales, on s'est attaché à prévoir tous les cas de figure, avec de multiples paramètres optionnels, si bien qu'implémenter ces normes signifiaient choisir entre toutes ces options. Était prévu en particulier le choix entre mode connecté, mode déconnecté (avec système de boîte aux lettres) et mode mixte. *Search & Retrieve* était un système idéal qui, pour être appliqué, nécessitait la définition et le test de profils : c'était lourd et complexe.

De son côté, l'ANSI (*American National Standards Institute*) a défini la norme Z39.50 plus simple. Et la montée en puissance d'Internet a bouleversé la donne, remettant en cause le modèle OSI avant de l'intégrer, et mettant fin au problème du mode de connexion (cela n'a aucune

importance de savoir si l'on travaille en mode connecté ou déconnecté). Les logiciels américains pour bibliothèque intègrent les uns après les autres la norme Z39.50, rejoints aujourd'hui par les produits européens.

Il s'agit donc de faire en sorte que Z39.50 et *Search & Retrieve* connaissent une évolution parallèle, manière pudique de dire que Z39.50 doit devenir une norme internationale. La chose est en bonne voie, avec la version 3 de Z39.50 (1995) dont le texte devrait servir de base à la nouvelle de *Search & Retrieve*, sans doute sous la référence ISO 23950. Les Américains acceptent cette internationalisation de leur norme, puisqu'ils accueillent des étrangers au sein du ZIG (Z39.50 Implementors Group), organisme chargé de gérer l'évolution de Z39.50.

Il faut enfin savoir que WAIS<sup>1</sup> (Wide Area Information Server) utilise la version initiale de la version 1 (1988) de Z39.50, qui n'a plus qu'un lointain rapport avec la version actuelle.

## **Z39.50, ou comment gérer les questions et les réponses indépendamment des systèmes**

Jean-Marc Czaplinski présente la norme Z39.50 en la définissant comme un protocole d'interrogation de bases de données (et pas seulement de catalogues de bibliothèque) reposant sur le mode client-serveur, sur un codage des données conformes aux normes de l'ISO (contrairement à WAIS) et sur le protocole TCP/IP, ce qui le rend compatible avec Internet. L'objectif est qu'un même logiciel client permette d'interroger n'importe quelle base : c'est un langage d'interrogation unifié indépendant des bases interrogées. Une même base peut être interrogée par des logiciels clients différents, qu'ils soient utilisés sur place ou à distance.

Z39.50 définit un ensemble normalisé de questions-réponses mais ne gère pas le traitement ni la présentation des données : c'est le logiciel client et le logiciel serveur qui s'en chargent. S'agissant de notices bibliographiques MARC, il prévoit leur fourniture en format d'échange (selon la norme ISO 2709<sup>2</sup>). La norme prévoit six syntaxes de recherche différentes, quatorze formats MARC et une centaine de champs bibliographiques (description bibliographique et des exemplaires), enfin une centaine de codes d'erreur.

À réception de la requête du client, le serveur établit et mémorise les réponses en un ensemble résultat ; il envoie un compte rendu d'exécution et un identifiant du résultat au client. Celui-ci peut alors demander tout ou partie du résultat. Ainsi, un logiciel client spécialisé Z39.50 peut-il interroger n'importe quelle base de données à condition qu'en aval ait été implémenté un logiciel serveur Z39.50 qui reçoit les questions et renvoie les réponses conformément à cette norme et fait l'interface avec la base cible.

Z39.50 prévoit des fonctions de contrôle d'accès (permettant par exemple de limiter l'accès de tels serveurs à des clients déterminés) et de contrôle de tarification (pour facturer le cas échéant les ressources utilisées sur le serveur). Enfin, sont prévus des services complémentaires (*extended services*) comme le prêt entre bibliothèques. C'est à la fois une richesse et un piège : la norme permet un grand nombre d'options et non un mode de fonctionnement unifié.

Divers types d'implémentation de Z39.50 sont possibles. Le client peut être situé sur le poste de l'interrogateur, et dialoguer avec un ou plusieurs serveurs distants, notamment via Internet. Mais il peut être également situé sur un antéserveur, donnant accès à une base ou un ensemble de bases, avec interfaçage permettant à l'interrogateur d'utiliser un simple navigateur World Wide Web.

Il ne faut pas confondre ce dispositif avec deux autres également mis en œuvre à

partir de ces navigateurs : les accès WAIS via un antéserveur ou l'interfaçage direct entre une base et World Wide Web.

## **SGML, ou que faire d'un langage de description logique de document ?**

François Role a présenté les perspectives ouvertes par la norme SGML (*Standard Generalized Mark-up Language* = langage standard généralisé de balisage). Née en 1986, cette norme est marquée par son origine éditoriale. Elle vise à décrire la structure logique d'un document (par exemple le découpage d'un texte en parties, chapitres, paragraphes, avec page de titre, table des matières, notes...). Il s'agit de définir une structure abstraite, permettant le cas échéant de produire un même document sous diverses formes (par exemple imprimé, CD-ROM, document en ligne).

SGML constitue un cadre général à partir duquel on peut définir diverses DTD (*Document Type Definition* = définition de type de document), par exemple pour les périodiques. Certaines DTD sont devenues des normes ou des standards<sup>3</sup>.

De grands éditeurs internationaux, comme Elsevier, ont investi dans SGML. Ils possèdent la maîtrise technique de ce langage, mais pour le moment ne s'en servent pas pour la diffusion des documents.

Il est difficile de déterminer si cet instrument de description de documents primaires peut devenir un outil pour les bibliothécaires. À tout hasard, la Bibliothèque du Congrès de Washington a réalisé une version SGML d'USMARC. Faut-il en conclure que SGML est appelé à remplacer les formats MARC ? La réponse de Christine Deschamps fut prudente. Une DTD peut contenir quelques codes à signification bibliographique, comme l'auteur, le titre. On pourrait donc déduire du document lui-même codé en SGML des éléments de description bibliographique.

1. Les outils WAIS permettent d'indexer simplement, par mot, des bases de données ou un ensemble de fichiers, bibliographiques ou non. Par extension, on qualifie de base « Wais » une base indexée grâce à ces outils.

2. La norme ISO 2709 pour l'échange de données bibliographiques permet notamment de structurer des notices de n'importe quel format MARC. Une grande partie des échanges bibliographiques s'effectue selon cette norme, par exemple à partir des CD-ROM bibliographiques.

3. HTML (Hyper text mark-up langage), langage de description des pages Web, est une évolution simplifiée de SGML, dont la DTD est adaptée à des besoins de présentation multimédia sur Internet.

## Besoins des bibliothèques et architecture client-serveur

Dominique Lahary a situé ces outils, normes et standards par rapport aux besoins des bibliothèques. En constatant tout d'abord que la profession se trouve confrontée à une avalanche de codes, sigles et acronymes qui, s'ils ne sont pas nouveaux (beaucoup sont apparus au milieu des années 1980), sont depuis peu brandis comme des impératifs catégoriques et affichés comme arguments de vente par les fournisseurs de logiciels. Comment s'y retrouver ? Le besoin d'information et de formation est immense. Nous avons mis dix ans à digérer UNIMARC. Et voilà que de nouveaux paysages apparaissent.

Un des rêves des bibliothécaires est l'accès à des catalogues et ressources hétérogènes, locales et distantes, à partir d'un

même poste, voire d'un même outil. Accès qu'on peut même souhaiter simultané à un ensemble de bases hétérogènes. C'est en quoi la norme Z39.50 peut les intéresser, et constituer une des réponses possibles.

La norme Z39.50 s'enrichit avec le temps : la version 2 comprenait 50 pages, et la version 3, 150. Sa complexité est surtout liée au fait qu'elle n'est pas utilisable telle quelle : son implémentation nécessite un certain nombre de choix, la définition de paramètres. Il y a là assurément un nouveau champ de normalisation, car pour faire dialoguer un outil standard avec des systèmes divers, c'est la recherche documentaire (critères, points d'accès, mode d'indexation, ...) elle-même qu'il faut normaliser. La norme Z39.50 le fait en partie, et l'on pourrait la définir sommairement ainsi : c'est du protocole plus des concepts bibliographiques. Cette norme définit en effet un certain nombre de notions, qui ne sont en définitive que les concepts bibliographiques classiques tels

que définis par les normes de catalogage et codés par les formats MARC (auteur principal, titre propre, etc.), mais en leur attribuant un jeu de valeurs propre à Z39.50, échappant ainsi à la diversité des étiquettes des différents formats MARC. Cette partie de la norme Z39.50, dit *jeu d'attributs bibliographiques n°1*, est directement lisible dans le texte par un bibliothécaire.

L'évolution de l'état de l'art, la généralisation de systèmes ouverts dans des applications client-serveur contribuent à concrétiser ce rêve des bibliothécaires. Il importe cependant de définir clairement de quoi il est question. Les concepts de client et de serveur sont des notions matérielles (tel poste de travail est client, tel ordinateur distant ou non est serveur), mais ce sont surtout des notions logiques : il y a des logiciels clients et des logiciels serveurs. Et l'on peut ainsi emboîter les pièces du puzzle suivant diverses configurations.