

FUTUR DES CATALOGUES CATALOGUES DU FUTUR

Actes de la journée d'étude du 21 juin 1996

Actes de la journée d'étude organisée à l'auditorium Vivienne de la BNF par l'Association des bibliothécaires français et l'Association des conservateurs de bibliothèque le 21 juin 1996, sous la coordination d'Antoine Provansal et Dominique Labary.

Vous trouverez plus haut dans ce même numéro trois des contributions qui ont été intégrées au dossier La bibliothèque en ligne.

Il s'agit de :

- « L'information secondaire au document primaire » par Catherine Lupovici ;
- « Z39.50 » par Jean-Marc Czaplinski et Yves Moret ;
- « Le jeu de puzzle de l'accès au catalogue » par Dominique Labary.

Neuf mois après

par Antoine Provansal

Conservateur, Bibliothèque nationale de France

Il y a trois sortes de bibliothécaires : ceux qui sont fascinés par leurs documents, ceux qui ne s'intéressent qu'à leurs lecteurs, ceux qui ont consacré leur vie au service du catalogue... Ceci n'est évidemment qu'une boutade mais qui montre à quel point les bibliothécaires définissent leur identité professionnelle à travers l'apprentissage du catalogage et de la bibliographie. Une collection ne devient une bibliothèque que si elle est organisée pour être consultée, donnée à voir à lire ou à entendre mais encore

aussi si elle est inventoriée, décrite, cataloguée.

François Dupuigrenet-Desroussilles le rappelle dans son intervention : la forme et la fonction des catalogues n'ont pas été bouleversées par l'apparition de l'imprimé. Inversement, on peut se demander si la numérisation croissante des données, révolution technologique qui se déroule sous nos yeux, va affecter les outils catalogographiques. C'est une des multiples questions posées lors de cette journée

dont le but était de confronter différentes approches disciplinaires, différentes perspectives, passées ou à venir, autour du thème « fédérateur » des catalogues. Partis de l'idée que le rapport entre information secondaire et information primaire¹ allait forcément évoluer avec l'apparition de nouvelles normes intégratrices (SGML,

1. L'information secondaire est celle que l'on trouve dans les catalogues et les bibliographies, l'information primaire désigne le contenu des publications elles-mêmes.

HTML, Z39.50), nous avons cherché à élargir le débat aux nouveaux services que peuvent offrir les catalogues et aux nouvelles demandes que l'on voit poindre en retour du côté des utilisateurs. Peut-on six mois plus tard reprendre les différents thèmes abordés et tenter de les replacer dans une réflexion plus large ?

Trois fils courent à travers ces interventions. Nous essaierons de repérer où ils se croisent et parfois se nouent. Le premier de ces fils est tendu par l'évolution des moyens de transmission du savoir. C'est l'histoire des médias, qui à travers l'apparition successive du manuscrit sur rouleau puis sur codex, du livre, de la presse, de l'informatique, détermine la façon dont les unités sont identifiées et cataloguées. On peut faire l'hypothèse que cette lente évolution historique va dans le sens d'une abstraction progressive des entités décrites. Un second fil suit les efforts normatifs internationaux. Rêvés par quelques visionnaires à la fin du XIX^e siècle et au début du XX^e, ils se sont mis en place d'abord à l'échelle des États-Unis, puis avec l'enthousiasme caractéristique de l'après Seconde Guerre mondiale, aboutissent aux ISBD, au format UNIMARC, puis à la mise en place des grands fichiers d'autorité. Cette information bibliographique qui est maintenant produite et diffusée à grande échelle par des fournisseurs publics ou privés, aura-t-elle encore un rôle important à jouer à l'ère de l'édition électronique ? Le troisième fil que nous suivrons est sociologique et politique, au sens premier de ce mot. Il part des recherches conduites sur l'interrogation des catalogues par les utilisateurs. Il montre à quel point la simplicité d'accès et de consultation des catalogues est une donnée essentielle pour permettre l'accès de tous, et singulièrement de ceux qui en sont exclus, à l'information et à la culture.

Les principaux outils bibliographiques se sont mis en place au Moyen Âge. Il s'agit de l'inventaire, qui répertorie les textes et les manuscrits présents dans un même lieu, des catalogues collectifs qui apparaissent dès les XII^e et XIII^e siècles en Angleterre puis en Allemagne, des bibliographies elles-mêmes qui s'affranchissent des collections particulières pour présenter les œuvres selon un classement « raisonné ». Lorsque l'imprimerie naît, les compagnons de Gutenberg et des maîtres rhénans se dispersent dans toute l'Europe

avec sur le dos leurs fontes ou leurs plaques xylographiques. Ils choisissent des villes commerçantes et libres. Ils datent leur production. Ces éléments forment l'adresse bibliographique de nos catalogues actuels : ville, raison commerciale, date. Les informations intellectuelles sont regroupées, dès le XVI^e siècle, sur un nouvel élément de présentation : la page de titre. Mais le principal changement qu'entraînent la diffusion et la conservation de l'imprimé, tient à la définition même de l'unité documentaire conservée et décrite. Les manuscrits étaient souvent des recueils factices reliés après coup, l'unité documentaire recherchée était donc l'œuvre, repérée par son incipit. Avec la reproduction imprimée, l'unité documentaire devient l'édition, qui est le produit d'une activité industrielle. Les catalogues décrivent alors ce niveau éditorial et non plus l'exemplaire possédé par la bibliothèque.

Avec l'apparition de la presse au XVII^e siècle, l'entité identifiée devient encore plus abstraite. Il s'agit d'un titre de périodique qui vit, meurt, renaît, change de nom et dont les bibliothèques possèdent rarement tous les fascicules. L'unité documentaire change une nouvelle fois avec les documents publiés (dans le sens de « rendu public ») sur Internet. Et se posent à nouveau, d'une façon qui est loin d'être résolue, les questions préalables au catalogage : de quelle version s'agit-il ? est-elle définitive ? quels en sont les responsables intellectuels ? surtout, quelles sont les limites du document, puisque des liens hypertextes le maillent dans un ensemble infiniment plus vaste... Des réponses pratiques doivent être rapidement données à ces questions si l'on veut que les documents électroniques soient catalogués et indexés et non simplement signalés par une fiche technique² et une adresse changeante, ce qui arrangerait bien les informaticiens. Pour les documents en ligne, un groupe de travail OCLC/NCSA³ propose une solution originale : chaque auteur d'un document ou plus exactement d'un objet documentaire (en anglais : *document like object*) pourrait remplir une notice descriptive comprenant les éléments essentiels d'identification : titre, responsabilités, numéro, etc.

2. C'est ce que l'on trouve, par exemple, dans Électre multimédia.

3. OCLC : Online Computer Library Center. NCSA : National Center for Supercomputing Applications.

Ces éléments, ou « *metadata* », seraient balisés par des marqueurs de type HTML ou équivalent. L'ensemble de ces éléments forme le *Dublin core set*, du nom de la ville de l'Ohio où se sont tenus les débats.

L'adoption de normes est préalable à toute entreprise bibliographique. Le fait de s'être mis d'accord au niveau international dans l'optique du contrôle bibliographique universel sur des règles de description communes et sur des formats compatibles et échangeables entre systèmes informatiques constitue une avancée considérable des trente dernières années.

Les normes internationales de description bibliographique (ISBD) développées par l'IFLA prennent en compte les besoins documentaires à un niveau de généralité suffisant pour être appliquées à des types de documents variés : textes, cartes ou partitions imprimés, audiovisuels, images fixes, documents électroniques. Ces ISBD rendent possibles les catalogues intégrés recensant tous les médias, catalogues que Francis Agostini propose d'appeler joliment « plurimédias ». De même, la guerre des MARC n'a pas eu lieu, puisque l'UNIMARC salvateur nous a permis depuis 1983 de récupérer ou de vendre nos notices sans perdre la richesse des formats internes particuliers avec lesquels il coexiste. La norme ISO 2709 a permis de consolider cet édifice, en spécifiant la disposition des données bibliographiques pour l'échange de notices MARC entre deux ordinateurs. Aujourd'hui, UNIMARC est la principale référence utilisée par les travaux internationaux dans le domaine bibliographique. Un projet européen comme UseMARCON, qui vise à permettre la conversion automatique de n'importe quel format MARC vers n'importe quel format MARC, utilise ainsi UNIMARC comme plate-forme.

Les deux nouveaux venus dans ce paysage normatif sont Z39.50 et SGML. La norme ANSI/NISO Z39.50, qui deviendra ISO 23950, s'appuie sur l'architecture client-serveur. Elle permet l'interrogation de bases de données hétérogènes et/ou distribuées. Par bases hétérogènes, on doit comprendre non seulement les bases bibliographiques en MARC, mais aussi dans les versions récentes de la norme, les bases documentaires qui ne sont pas en MARC. La norme Z39.50 permet de ne pas tenir compte de la syntaxe d'interro-

gation propre à chaque base, ni du système qui les supporte, c'est ce dernier aspect que l'on appelle interopérabilité entre systèmes. Elle permet de formaliser les requêtes lancées sur ces bases et la transmission des réponses. Des profils sont développés pour des besoins particuliers, comme le profil CCFr qui sera mis en œuvre dans le système d'information du Catalogue collectif de France. Ce profil est élaboré conjointement par la Bibliothèque nationale de France, le MENESR⁴ et la mission pour le CCF.

Les normes de balisage logique des documents, SGML pour le texte, HTML⁵ pour le multimédia, structurent les données textuelles ou audiovisuelles le plus tôt possible dans la chaîne de production des publications. Développées à l'origine par l'*American Publishers Association*, SGML permet le repérage dans un texte avant publication de tous ses éléments formels, des éléments bibliographiques traditionnels (titre, auteurs, éditeur commercial, nombre de pages, etc.) mais aussi d'éléments que ne donnent pas les catalogues traditionnels mais qui intéressent fortement les chercheurs (sommaire, titres de chapitre, notes de bas de page, etc.).

Si la normalisation Z39.50 permet d'ignorer superbement les formats des bases

4. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.

5. Respectivement Standard Generalized Markup Language et Hypertext Generalized Markup Language.

documentaires interrogées, SGML et les normes associées permettent, quant à elles, de ne plus séparer l'information publiée en deux blocs étanches : le texte intégral d'un côté, les références documentaires de l'autre.

L'intégration du signalement du document et du texte numérisé est maintenant un horizon proche de la bibliographie. Ce qui apparaît déjà dans des réalisations aussi séduisantes que le catalogue de la bibliothèque de Valenciennes, c'est la possibilité d'afficher des « bouts » du document, tel extrait vidéo, tel frontispice, tel sommaire de périodique. Reste à savoir ce qui intéresse le plus les usagers qui interrogent nos catalogues, en fonction des types de demandes répertoriés : loisir ou distraction, demande scolaire ou universitaire, recherche. Notons qu'Hervé Le Crosnier a des idées très précises sur la question ! Ainsi, de même que la pression des utilisateurs a permis la généralisation des interfaces de type Windows pour l'interrogation des systèmes informatiques, de même, l'habitude d'Internet rendra difficilement supportable l'absence de liens hypertextes entre notice catalographique et document décrit.

Les pratiques réelles d'interrogation et de consultation des utilisateurs forment un autre aspect essentiel des recherches sur les catalogues. Les programmes de recherche de type Parinfo et l'ensemble des travaux sur les catalogues interactifs ont permis de mieux comprendre ce qui se passait entre l'homme qui interroge et

la machine qui lui répond (pas toujours malheureusement !).

Les recherches sur les OPAC⁶ et l'interrogation des bases de données permettent d'entrevoir la distance entre les intentions des concepteurs d'un système et la pratique de ceux qui l'utilisent. Ainsi, le développement de gigantesques systèmes documentaires qui intègrent les fonctions les plus diverses ou de normes internationales qui mobilisent des armées d'experts ne garantissent pas la réussite des projets. La définition de l'ergonomie de la consultation, qui est heureusement une des premières étapes de la mise en place du système d'information de la Bibliothèque nationale de France, est un élément clé de la réussite des catalogues du futur. Cela ne suffit pas. L'ouverture des bibliothèques à tous passe par la prise en compte des indices de satisfaction des utilisateurs, de leurs échecs, de leur plus ou moins grande familiarité avec l'univers scolaire, universitaire ou documentaire. N'étant pas soumises à la pression du marché, les bibliothèques peuvent définir leur priorité face à une demande multiforme et envahissante. Seules leur volonté et leur éthique professionnelle les conduiront à faire des catalogues des outils simples, voire attractifs, accessibles à tous et tournés vers l'obtention la plus rapide du document sous sa forme papier ou numérique. C'est la condition d'une démocratisation de leur accès.

6. Online Public Access Catalog.