

# LES BIBLIOTHÈQUES ET LE DÉFI DE L'ÉDITION ÉLECTRONIQUE

L'année 1995 aura été une année particulièrement importante dans l'avènement de l'édition électronique, la vraie. Non pas la numérisation de l'édition papier qui est une conversion rétrospective, mais l'édition électronique d'origine, dans une forme qui n'a plus pour unique objectif de produire un document imprimé ou son équivalent image sur CD-ROM.

Chez les grands éditeurs STM (sciences, technologie et médecine), cette année est marquée par l'abandon (ou presque) du discours sur le fantastique avenir de l'édition électronique, auquel s'attachaient à la fois une attitude prudente sinon peureuse et une agressivité mêlée d'une volonté impérialiste vis-à-vis des professions de la documentation. Elle représente le début du passage à la réalisation. Les éditeurs semblent rassurés sur leur rôle et leur positionnement futur. Elsevier a engrangé les premiers résultats de ses études marketing et techniques dans le cadre du projet TULIP et étend ses projets à l'Europe. Le projet ELSA (Electronic Library SGML Application), qui a commencé au début de l'année, a déjà fait progresser les solutions techniques, mais également la réflexion sur la nature même de l'information, ses relations avec sa structure et avec sa présentation.

Chez les bibliothécaires, 1995 représente une charnière entre l'intérêt pour l'idée de bibliothèque électronique et la nécessité technique de se préparer à la gestion des documents électroniques, dont on sait désormais l'échéance proche. Les thèmes déve-

loppés lors des conférences de l'IFLA montrent bien ce passage et la conférence d'Istanbul est la première à avoir traité du document électronique en tant que phénomène concret, contemporain, irréversible, montrant à l'évidence l'implication des bibliothèques, plus encore qu'il y a vingt ans pour les documents audiovisuels.

Depuis les déclarations du sénateur Al Gore en 1992 sur les autoroutes de l'information jusqu'à la réunion du G7 en février 1995, la prise de conscience politique de l'importance économique et stratégique de l'information, dans un monde de télécommunications, a cheminé peu à peu.

Par-delà les visions de bibliothèques virtuelles qui ouvrirent la voie, les bibliothécaires ont maintenant conscience d'être au seuil d'une révolution des pratiques du métier qui, bien que certaine, ne permet pas encore de discerner ses contours précis ni ses implications juridiques ou économiques.

## **L'économie du document papier**

Le document papier est un produit manufacturé conçu par un éditeur, réalisé par des auteurs pour le contenu et par un imprimeur pour la forme. Ce produit est achevé à ce stade. Il reste à le distribuer et le vendre. Les composants de sa fabrication (papier, brochage, couverture...) ainsi que son transport coûtent cher. Il est fabriqué en une certaine quantité, *a priori*, en fonction des prévisions de vente. Enfin

CHRISTIAN LUPOVICI

Institut de l'information scientifique et technique

E-mail : lupovici@inist.fr

c'est un produit qui se vend et l'acquéreur en devient propriétaire, même si subsistent des droits attachés à la propriété intellectuelle et à l'économie de sa reproduction.

### **Les métiers traditionnels de la chaîne de l'information**

C'est autour de cet objet que se sont constitués des métiers tout au long de la chaîne de l'information, depuis l'auteur, l'éditeur et l'imprimeur, jusqu'aux distributeurs intermédiaires que sont les librairies et agences d'abonnements, les bibliothèques et centres de documentation, pour aboutir à l'utilisateur final qui a besoin d'une information.

Chaque intermédiaire, qui doit traiter le document pour en gérer sa part transitoire, a développé un savoir-faire fondé sur les caractéristiques physiques du document. Le document papier, mais il en va de même du disque ou de la cassette, est traité comme un objet dont on extrait des éléments pour lui donner, de façon externe, un accès physique (classification, indexation matière, indexation auteur et titre).

Dans cette chaîne, chaque maillon a défini son propre métier en fonction de l'objet à traiter. Pour ne s'en tenir qu'au métier des bibliothèques, sont nées des spécialités comme le catalogue, l'indexation, le résumé documentaire, la classification, l'interrogation documentaire, etc.

Chacune de ces techniques a fait l'objet d'une normalisation de ses procédures pour garantir à l'utilisateur une meilleure efficacité dans sa recherche et une plus grande compatibilité des données au niveau national et international.

Ce qui est clair, dans cette économie, c'est que chaque maillon dans l'accès à l'information a développé ses contraintes propres en fonction du type physique du document à traiter et du support d'information utilisé (fiches, OPAC, niveau de sophistication de l'interrogation...). Chacun a développé sa technique dans son environnement, indépendamment des autres. Il y a le monde des auteurs, celui des éditeurs, celui des

producteurs de bases de données et celui des bibliothèques. Chacun de ces mondes a sa propre filière de normalisation, qui donne de la cohérence à son environnement, mais en se coupant des autres mondes.

### **L'économie du document électronique**

L'édition électronique bouleverse l'économie du système de fabrication, de distribution et de recherche des documents. Du point de vue technique, le document électronique est davantage le véhicule d'un contenu que sa matérialisation, qui est une contrainte et non plus le but ultime.

Le document perd ses repères physiques. Le livre est une collection de chapitres et le périodique une collection d'articles. L'article lui-même est un assemblage d'éléments multimédias, repérables et manipulables séparément, dans le cas de documents de type SGML (Standard Generalized Markup Language).

© INIST

Le document n'est donc plus seulement lisible séquentiellement ou dans un contexte fixe de mise en pages. Il peut être accessible par élément ou ensemble d'éléments sélectionnés selon leurs caractéristiques physiques (les schémas ou les séquences d'images ou de son) ou selon leurs caractéristiques logiques (introduction, résumé, conclusion, bibliographie...).

### **Transfert des savoir-faire d'un métier à l'autre**

Même si les métiers actuels conservent leur place relative dans la chaîne de l'information, le document électronique resserre leurs activités à l'essentiel de leur valeur ajoutée dans le traitement de l'information, qui est leur seule justification.

Pour ne s'en tenir qu'aux évolutions qui concernent directement les

tion du résumé devient automatique, dans le même format d'exploitation des données bibliographiques et du texte intégral, comme le préfigure le service CAPCAS d'Elsevier Science, tel qu'il est déjà utilisé en production à l'INIST.

### ***L'indexation***

Le document électronique contient en lui-même tous les éléments d'information qui servent à son indexation et, à défaut du texte intégral, le résumé joint au titre. Le document n'est plus une source externe que l'on analyse pour saisir les mots-clés qui serviront d'accès au texte, il est en soi la source directe de l'information indexée. Il peut donc servir de base à une indexation automatique, effectuée de façon dynamique par un logiciel d'indexation du texte intégral.

Cette technique permet non seulement une plus grande homogénéité de l'indexation, mais, n'étant plus liée à la notice bibliographique, la sophistication de la technique de filtrage peut évoluer dans le temps, au fur et à mesure que les logiciels de traitement de l'information évoluent et ceci sans modifier l'enregistrement source.

Si l'indexation manuelle est appelée à disparaître, le besoin de performance des traitements linguistiques, tant au niveau de la question en langage naturel qu'à celui de la traduction automatique des termes indexés, exige l'adjonction de bases de connaissances multilingues. Dans ce cas, il y a un enrichissement du travail documentaire par rapport au thésaurus classique.

### ***La classification***

La classification, comme moyen de trier l'information par thème, reste un problème à résoudre en aval de l'éditeur, en tout cas pour des exploitations spécifiques, car on peut imaginer que l'éditeur indique, à l'origine, les codes de classement généralement utilisés par ses clients (Library of Congress Classification, National Library of Medicine, Dewey). La classification, comme rangement des

© INIST

bibliothèques et centres de documentation, certaines fonctions sont transférées en amont, vers l'éditeur, d'autres, nouvelles, lui échoient.

### ***Le catalogage***

Le catalogage partagé avec la dérivation des notices bibliographiques de catalogues coopératifs, tels que OCLC, RLIN ou PICA, avait déjà relativisé la fonction de catalogage dans les activités de bibliothèque.

Aujourd'hui, avec le document électronique qui comporte en lui-même les données catalographiques, nous pouvons annoncer l'éradication du

catalogage dans les bibliothèques pour tous les documents électroniques. Néanmoins, la compétence documentaire des bibliothécaires et documentalistes reste indispensable dans le dialogue nécessaire avec les éditeurs, qui sont désormais les catalogueurs de leur production.

### ***Le résumé documentaire***

Le résumé documentaire n'est plus effectué, depuis longtemps, en interne, par les grands producteurs de bases de données. Ils prennent le résumé d'auteur, saisi à partir de l'édition papier, par reconnaissance optique de caractères (OCR). Avec le document électronique, la récupéra-

documents, n'a évidemment plus aucun sens dans un environnement électronique.

### ***Le stockage***

Les bibliothèques ont été conçues pour gérer et stocker différents formats physiques de documents et le format a souvent déterminé la localisation du document pour une optimisation de l'espace. Dans le cas de l'édition électronique, en l'état actuel des choses, ceci est transposable. En effet, si les grands éditeurs estiment pouvoir fournir aux centres documentaires et réseaux locaux d'université des documents en SGML – encore existe-t-il plusieurs DTD (définition de type de documents) – d'autres formats sont déjà distribués, tels PDF (Portable Document Format), HTML (HyperText Markup Language), Postscript, ASCII et des formats image TIFF (Tagged Image File Format) ou GIF (General Information File)...

C'est l'intermédiaire, la bibliothèque, qui doit pouvoir gérer cette disparité, pour offrir à son utilisateur final le format qui lui est le mieux adapté. Ce sera dans la plupart des cas une forme papier. Ainsi les bibliothèques devront acquérir une compétence d'impression comparable à celle des imprimeurs.

### ***Le nouveau rôle des bibliothèques***

Avec le document électronique et les autoroutes de l'information, les bibliothèques vont abandonner progressivement les tâches techniques, catalogage, bulletinage, indexation, classification et même interrogation des banques de données, qui furent une grande partie de leurs activités. Ces activités seront effectuées ailleurs, ou bien ne se justifieront plus tout à fait face à la simplification des interfaces utilisateurs. Néanmoins, la compétence (surtout théorique) dans ces domaines restera (avec sans doute plus d'acuité) indispensable et partie intégrante du métier de bibliothécaire/documentaliste.

En effet, le catalogueur de l'avenir n'est pas le spécialiste de la ponctua-

tion ISBD. Il ne sera plus non plus le spécialiste des formats Marc, mais l'expert des répertoires de données associées aux applications mises en œuvre et un spécialiste des DTD bibliographiques et des documents en SGML. Il devra être capable de discuter de l'évolution et de la qualité du format avec l'éditeur qui le produit et les gestionnaires du système d'information qui les exploitent. Il devra savoir quels éléments sont nécessaires pour produire les services à l'utilisateur final.

## ***LE RÔLE DE GUIDE, DE CONSEIL, DE FORMATEUR À L'UTILISATION DES OUTILS ET DES RÉSEAUX DEVIENT PRÉPONDÉRANT***

Le catalogueur matière ou l'indexeur de notices bibliographiques deviendra un expert du traitement du langage et de la constitution de ressources terminologiques, et en particulier, un spécialiste de bases de connaissances.

Ces transformations, loin d'être réductrices, sont au contraire extrêmement valorisantes, mais elles réclament moins d'emploi pour plus de qualification.

Les tâches techniques (y compris le magasinage) vont se réduire peu à peu mais très sensiblement. Une réflexion sur le métier conduit à estimer que l'essentiel de l'expertise des bibliothécaires consistera en une fonction de filtre et de guide. L'abondance des sources d'information et des moyens d'y accéder ont généré le paradoxe qu'il est difficile de trouver la bonne voie vers l'information pertinente et que l'utilisateur final, qui n'est pas, et n'a pas

à être un spécialiste de l'information, perd trop de temps (donc d'argent).

Le rôle de guide, de conseil, de formateur à l'utilisation des outils et des réseaux devient donc prépondérant. C'est en somme l'évolution logique du rôle du bibliographe et du *reference librarian*, avec ce qu'il suppose de professionnalisme mais aussi de rapport privilégié avec l'utilisateur. Pour leur documentation, les entreprises et bientôt les chercheurs du secteur public seront prêts à soustraire la fonction documentaire auprès de ceux qui savent l'assurer de façon plus efficace.

La bibliothèque aura également pour rôle de préparer des produits à plus forte valeur ajoutée. Des produits et services personnalisés, automatiques, sont nécessaires, dans un cadre contractuel, et prévoyant un engagement sur la qualité, fondé sur un travail en commun de définition des besoins.

Il s'agira de fabriquer des profils de veille documentaire qui, dans leur principe, peuvent rester relativement standards, mais qui seront obtenus en mettant en œuvre des techniques beaucoup plus élaborées qu'une simple équation de recherche documentaire.

Ces produits pourront être aussi très ciblés pour certains : détection des « signaux faibles » d'une technologie émergente ou de la stratégie de la concurrence, et pourquoi pas, fabrication des dossiers de synthèse ou d'alerte sur les sujets les plus sensibles pour la population desservie. Il est également évident que ces produits ne doivent pas reposer seulement sur le dépouillement de la littérature conventionnelle, mais aussi sur la littérature grise, en particulier les sources électroniques repérées sur les réseaux, dont il faut cependant relativiser le degré de fiabilité.

Tous ces éléments existent déjà et l'évolution prévisible n'est pas une révolution, mais plutôt une systématisation de l'emploi de ces nouvelles technologies de l'information et de ces nouveaux services à valeur ajoutée, dans des milieux professionnels qui ne les pratiquaient pas jusque-là.

Enfin, la bibliothèque comme centre de concentration de moyens, d'infrastructure informatique en particulier, sera le lieu privilégié de concentration des moyens de visualisation ou d'impression de documents sophistiqués. Mais on peut également imaginer que la bibliothèque ne soit plus identifiable à un ou plusieurs lieux, mais un service dont la localisation peut être très éclatée. Il n'empêche que la relation étroite avec l'utilisateur final est la part non réductible du métier de bibliothécaire.

© INIST

### **L'environnement juridique**

Nous connaissons bien maintenant le cadre juridique du document édité sous une forme papier. Le document électronique échappe quelque peu à ce cadre. Bob Wedgworth a bien montré, dans son discours d'ouverture de l'IFLA 95 (10), que les éditeurs imposent, pour leurs documents électroniques, non plus des contrats de vente de document, en laissant à la loi le soin de régler le cadre d'application du droit de copie, mais des contrats de licence d'utilisation de l'information.

Le document électronique peut donc être considéré comme n'étant pas possédé par la bibliothèque (qui le conserve, le prête, etc.), mais en dépôt temporaire pour une exploitation réglementée en fonction de contraintes techniques et économiques négociées.

Le cadre juridique qui se dessine pour le document électronique ressemble

plus à celui des logiciels ou des banques de données qu'à celui traditionnel du livre. Une licence est délivrée pour l'exploitation contrôlée d'un document donné, pour un site donné. La rémunération s'effectue à la transaction dans le cas de la fourniture de documents ou par abonnement dans le cas de réseau local, en fonction du nombre d'utilisateurs ou de postes de travail ou d'accès simultanés.

### **L'INIST : vers la bibliothèque virtuelle**

Le thème de la bibliothèque virtuelle a été traité dès 1992 par ELAG (European Library Automation Group). Néanmoins, les applications mettant en œuvre ce concept n'existent toujours pas de façon significative. Il est vrai qu'il faut pour cela disposer d'une palette importante de ressources électroniques, ce qui commence à peine à devenir une réalité

avec le Web à travers Internet.

L'INIST, pour sa part, construit sa stratégie d'avenir en déclinant les conséquences de l'évolution des technologies, non seulement sur les métiers, mais également sur l'attitude des clients actuels et potentiels.

L'INIST a été fondé sur une juxtaposition des métiers traditionnels de la documentation et selon une organisation de la production qui suit la chronologie de la chaîne documentaire, depuis les acquisitions jusqu'à la mise en ligne des bases de données et la fourniture des documents.

L'acquisition de documents électroniques, l'utilisation des techniques de traitement automatique de l'indexation sur le texte intégral et les possibilités de transfert électronique des documents impliquent forcément une évolution profonde des métiers et de l'organisation du travail.

Si la forte spécialisation scientifique des documentalistes, qui analysent les documents pour les bases Pascal et Francis, pouvait être jugée hors de propos dans le contexte d'une indexation manuelle, elle devient un atout prépondérant dans le contexte d'une indexation automatique. En effet, il devient tout à fait pertinent de disposer de spécialistes des domaines scientifiques pour construire les bases de connaissances nécessaires à l'indexation automatique, à la recherche en langage naturel et à la traduction automatique. C'est ainsi que va évoluer le métier d'indexeur.

La base Pascal, qui avait été conçue essentiellement pour une interroga-

tion bibliographique classique, doit désormais servir également de réservoir source pour la veille technologique des entreprises. L'INIST a dû modifier la couverture des documents dépouillés et ajoute désormais les éléments qui pouvaient faire défaut (complétude des résumés, des auteurs et des affiliations), anticipant ainsi sur les données qui proviendront bientôt de l'édition électronique.

Par ailleurs, si le chercheur est assez bien outillé pour trouver l'information qu'il cherche avec les réseaux de communication actuels, le public de l'entreprise (et *a fortiori* le grand public) est démuné face à la complexité de l'accès à l'information scientifique et technique (IST). Il est souvent noyé sous une masse d'informations dont peu sont pertinentes ; de plus, le vocabulaire documentaire employé (souvent en anglais) lui est étranger. Outre les modifications nécessaires au traitement des sources d'information, l'INIST se devait de concevoir un service complet d'accès à l'information pour sa clientèle cible.

## Agades

En déclinant les conséquences des attentes de son public, l'INIST a donc élaboré un projet de service d'information qui a été proposé au ministère de l'Industrie dans le cadre des travaux consacrés aux autoroutes de l'information. Le projet AGADES (Accès général à la documentation électronique spécialisée) repose sur une série d'éléments qui sont pour la plupart à l'état de test, mais qui n'ont pas encore été assemblés. Ce projet rassemble bon nombre des grands acteurs de l'IST en France, comme Questel-Orbit, ERLI, Jouve SI, l'Institut national de la propriété industrielle, la Chambre de commerce et d'industrie de Paris, de grandes entreprises, ainsi que la Bibliothèque nationale de France et des bibliothèques universitaires.

– Le premier élément est un nouveau concept de la base Pascal, considérée comme une base de première approche, essentiellement riche en sciences appliquées plutôt que fondamentales, permettant la navigation et

donc l'affinement de la question.

– Le second élément concerne l'interrogation : l'effort est mis sur l'accès en langage naturel, en français. Ceci suppose non seulement le traitement linguistique de la question, mais encore la traduction simultanée vers l'anglais pour assurer l'exhaustivité de la recherche.

– Le troisième élément concerne la fourniture électronique de documents : l'utilisateur pourra commander et recevoir immédiatement sur son poste, dans la même session que l'interrogation, le document signalé, s'il est sous forme électronique.

– Le quatrième élément est le reroutage de la question ou de la commande du document vers des sources externes identifiées dans un catalogue des ressources. Celles-ci peuvent être des bases de données bibliographiques ou factuelles, ou des localisations de documents dans des bibliothèques utilisant un système de transfert électronique des documents (par exemple, la station développée dans le projet EDIL-Electronic Document Interchange between Libraries).

## La possession n'est plus un impératif

Le document électronique abolit l'espace entre sa localisation et l'utilisateur. Le réseau Internet montre la puissance de ce concept. Dès lors, la possession du document primaire n'est plus un impératif d'efficacité. L'acquisition doit être mise en balance avec la fourniture à distance, à la demande. La politique de développement des collections est une nécessité qui intègre désormais l'économie globale de l'accès au document.

L'INIST, en mettant en place une infrastructure qui permet une interactivité entre l'utilisateur et la source de l'information, établit les bases d'une réelle bibliothèque virtuelle où des entreprises pourront sous-traiter leur documentation comme elles le font du ménage ou de la cantine. Il y a fort à parier que, dans les cinq ans à venir, se dessinera un réseau de bibliothèques, dont les fonctions élec-

troniques devenues prépondérantes, réaliseront la bibliothèque virtuelle du siècle prochain.

Octobre 1995

## BIBLIOGRAPHIE

**1. Gronk, Ronald D.** – « Unlocking data's content », *Byte*, vol. 18, n° 10, Sept. 1993, p 111-120.

**2. Hunter, Judy F. ; Cotter, Gladys A.** – *The Virtual Library : Coming of Age*, AGARD meeting on international high speed networks for scientific and technical information, Oct. 1993, p 5/1-5/5.

**3. Long Maurice.** – « The future of scientific communication between authors and users : how will publishers and librarians fit in ? », *Knowledge for Europe : Librarians and Publishers Working Together / European Conference*, 11-13 Nov. 1992, Brussels. – Bruxelles : Saur, 1993, p 130-143.

**4. Richardson, John.** – « Pointing the indexfinger », *Information Management and Technology*, vol. 27, n° 1, 1994, p. 18-20.

**5. Role, François.** – « De la lettre au sens : les recherches en texte intégral », *Documentaliste-Sciences de l'Information*, vol. 30, n° 3, 1993, p 136-146.

**6. Schauder, Don.** – « Electronic publishing of professional articles : attitudes of academics and implications for the scholarly communication industry », *JASIS*, vol. 45, n° 2, 1994, p 73-100.

**7. Siddigni, Moid A.** – « Full-text databases », *Online Review*, vol. 15, n° 6, 1991, p 367-372.

**8. Van Marle, Gérard A. J. S.** – « Electronic serial publishing and its effects on the traditional information chain », *Serials*, vol. 7, n° 1, March 1994.

**9. Woodward, Hazel.** – « The impact of electronic information on serials collection management », *IFLA Journal*, vol. 20, n° 1, 1994, p 35-45.

**10. Wedgeworth, Robert.** – « Beyond the limits of space and time », 61<sup>st</sup> Conference générale de l'IFLA, 20-26 août 1995 Istanbul, Turquie.