

# Chapitre 7

## Services d'information, bibliothèques et archives

OLE HARBO

Professeur à l'École royale de bibliothéconomie  
et de sciences de l'information,  
Copenhague (Danemark)

### UN DOMAINE LARGE ET DIVERSIFIÉ

On trouve dans ce domaine des établissements très divers. Les archives et les bibliothèques dépendent généralement des instances fédérales, de celles des États ou des collectivités locales qui en sont les propriétaires et assurent leur fonctionnement — ou, du moins, leur apportent un soutien vigoureux. Les services d'information, définis ici comme des entreprises distinctes ou des éléments d'une grande organisation, relèvent généralement du secteur privé. Toutefois, certaines de ces organisations sont des entités publiques ou semi-publiques fonctionnant selon les lois du marché. La taille de ces établissements varie aussi beaucoup : ce peut être des entreprises dont une seule personne est propriétaire et assure le fonctionnement, de petites bibliothèques publiques ou scolaires locales ayant des collections minuscules et n'employant qu'une personne à temps partiel, mais aussi de grands producteurs et vendeurs multinationaux de services d'information, ou des archives et des bibliothèques nationales massives détenant des millions de documents et employant des centaines de spécialistes. Les facteurs géographiques et économiques accroissent encore la diversité de ces établissements. Ces différences se répercutent bien évidemment sur l'utilisation plus ou moins intensive et sophistiquée des technologies modernes de l'information et de la communication.

Plusieurs aspects de l'utilisation de ces technologies dans les secteurs des archives, des bibliothèques et des services d'information, où elles sont déjà bien implantées, seront examinés dans le présent chapitre. Cette approche permet d'étudier des problèmes qui se posent dès à présent dans un nombre d'établissements qui est déjà important et risque encore d'augmenter considérablement dans l'avenir proche. Les questions abordées seront d'ordre professionnel et institutionnel plutôt que technique. Le but est de mettre en évidence l'impact des pratiques et des principes professionnels sur une grande majorité d'établissements fournissant des services d'information et sur leurs utilisateurs.

## IMPACT SUR LES SERVICES

Le manque d'espace résultant de la croissance des collections est un problème aussi ancien que le secteur lui-même. Il se pose non seulement pour les archives et l'imprimé, mais aussi pour divers types d'enregistrements sonores, de films, de vidéos, etc. On a résolu ou tenté de le résoudre en rangeant les collections de manière à occuper moins de place, en les élaguant ou en utilisant des procédés tels que le microfilmage pour réduire la taille des documents. L'informatique a ouvert d'immenses perspectives de réduction du fait de la capacité de stockage des ordinateurs qui peut être effectué dans les établissements eux-mêmes ou à l'extérieur.

La vitesse avec laquelle de grandes quantités de données peuvent être manipulées est une autre caractéristique de l'informatique. Cela permet de retrouver quelques documents intéressant un utilisateur parmi les milliers de ceux qui portent sur le même sujet. L'efficacité des canaux de communication assure un accès très rapide au contenu des documents, où qu'ils se trouvent et quels que soient les heures d'ouverture des établissements et les fuseaux horaires. Les problèmes de traitement de la quantité importante et croissante d'information figurant sur l'Internet seront examinés plus loin.

L'expression « secteur des services d'information » sera utilisée pour l'ensemble du domaine : archives, bibliothèques et services d'information. Il est bien évident que ce secteur se prête à l'utilisation des technologies de l'information et de la communication mais il ne faut pas perdre de vue que celles-ci sont soumises aux lois du marché. C'est un des facteurs qui entraînent les déséquilibres bien connus de l'utilisation de ces technologies par les pays aussi bien développés qu'en développement, et à l'intérieur d'un même pays, entre les secteurs de la défense, de la recherche spatiale, des affaires, etc., d'une part, et ceux qui se rapportent aux activités artistiques et de loisirs, d'autre part. Ce type de développement est une des

caractéristiques de ce domaine depuis que les ordinateurs sont apparus peu après la fin de la seconde guerre mondiale.

Leur orientation de plus en plus nette vers une utilisation **individuelle** au foyer ou sur le lieu de travail est une autre caractéristique importante des technologies de l'information et de la communication. C'est à bien des égards l'inverse de ce qui se passe dans les établissements consacrés à l'information où les services sont fondamentalement **collectifs** et où le même livre est lu par de nombreuses personnes. L'individualisation est manifeste pour le téléphone, à partir d'appareils traditionnels ou portables, et les services connexes tels que la télécopie. Il en va de même pour la télévision et les services d'information qui y sont liés tels que le télétexte, pour les enregistrements sonores et pour les vidéos. De plus, l'ordinateur est lui-même maintenant devenu **personnel** avec la multiplication des PC. Ces technologies et ces produits orientés vers l'individu peuvent tous être reliés grâce à des technologies numériques et sont utilisables partout où des signaux peuvent être transmis.

Il y a toujours eu au cours de l'histoire — et c'est encore le cas aujourd'hui — des périodes de transition plus ou moins courtes ou plus ou moins longues pendant lesquelles les technologies destinées à l'individu sont utilisées collectivement. Même dans les pays où les téléphones sont nombreux, il existe des cabines téléphoniques destinées à un usage collectif et il est encore possible de regarder un match de football à la télévision avec des amis et des étrangers dans un bar. Les utilisations collective et individuelle de livres coexistent depuis des siècles, et celle des productions audiovisuelles apparues plus récemment depuis des décennies. Dans les pays où le recours aux technologies individuelles est moins développé, une utilisation aussi collective que possible optimise évidemment l'emploi de ressources limitées, mais peut, comme on le verra plus loin, poser certains problèmes. Il convient aussi de préciser que certains types d'utilisation collective, par exemple dans les bibliothèques publiques,

sont un élément de la fonction sociale de ces institutions. Ils doivent donc être maintenus quel que soit le degré d'individualisation de la technologie. Comme dans toutes les institutions et entreprises, les technologies orientées vers l'individu sont utilisées par le personnel du secteur des services d'information pour la communication interne et externe. Ces aspects ne feront pas ici l'objet d'une analyse plus détaillée. Ce sont essentiellement les technologies numériques utilisées en informatique et les techniques connexes de communication employées pour les activités propres au secteur des services d'information qui seront examinées.

## GRATUIT OU PAYANT

Le paiement éventuel des services fournis par les bibliothèques fait depuis longtemps l'objet de discussions. Les fournisseurs privés d'information sous forme de documents ou de services n'ont pas de tels débats puisque ce sont les marchés qui déterminent les prix. Deux positions idéologiques s'affrontent en ce qui concerne le paiement des services de bibliothèque : les uns privilégient un accès égal et facile à des services gratuits, les autres considèrent que les mécanismes de régulation par le marché plaident en faveur du paiement. L'accès à des services gratuits est généralement l'option jugée la meilleure par les bibliothécaires et les politiciens.

Pour ce qui est des technologies informatiques, la question du paiement a deux autres aspects. Premièrement, il est possible de mesurer l'utilisation individuelle des services électroniques et il est donc beaucoup plus facile de fixer des prix que pour les services de bibliothèque traditionnels, qui sont essentiellement collectifs. Deuxièmement, les bibliothèques sont confrontées à des problèmes économiques. Elles doivent continuer à mettre livres et périodiques à la disposition des lecteurs tout en acquérant du matériel informatique et en réglant les factures inhérentes aux services électroniques et à l'utilisation des systèmes de

télécommunication. Ces évolutions ont, dans de nombreux pays, nourri le débat relatif au paiement par les utilisateurs et il a été suggéré que l'on pourrait faire une distinction entre les services de base et les services à valeur ajoutée, ou même entre l'information imprimée et électronique. Le coût élevé de la mise au point de matériels et de logiciels, et du développement de la collecte de données a créé des monopoles très puissants ou des conditions analogues à celles d'un monopole. De ce fait, les prix sont élevés et posent des problèmes de paiement aux établissements des pays en développement, mais aussi à la plupart des bibliothèques des autres régions du monde.

On commence aujourd'hui à s'organiser pour réagir. Ainsi, le personnel des universités, des bibliothèques et des archives, qui sont d'importants fournisseurs d'information pour les revues scientifiques, les recueils de résumés, etc., s'attache de plus en plus à produire lui-même, par exemple, des revues électroniques afin d'en réduire le coût. L'initiative anglaise de licence de sites pilotes (Pilot Site Licence Initiative, 1996-1998 [PSLI]), consortium créé pour arrêter la hausse des prix des périodiques, en est un exemple. Depuis janvier 1999, cette action se poursuit pour trois nouvelles années sous la forme d'un nouveau projet, l'initiative nationale de licence de sites électroniques (National Electronic Site Licencing Initiative [NESLI]) qui a pour objectif de regrouper tous les établissements d'enseignement supérieur du Royaume-Uni en un consortium. Après un appel d'offres de l'Union européenne, l'entreprise Swets et Zeitlinger a été désignée comme agent pour l'interface commune et les négociations avec les éditeurs.

## L'IMPACT SUR LES TECHNIQUES

### Préservation et stockage

Il est d'autant plus difficile de prendre les mesures nécessaires pour préserver les différents supports de stockage (papier, bandes magnétiques, vidéo ou disques) que, sauf pour le papier, leur durabilité est

## Encadré 7.1 → Le développement des nouvelles technologies de l'information dans les bibliothèques

### L'informatisation des bibliothèques

Les bibliothèques ont très tôt engagé l'informatisation de leur gestion. Depuis l'invention du format MARC (Machine Readable Cataloguing) au milieu des années 60, de nombreux pays ont adopté une variante nationale de ce format. Actuellement, les bibliothèques des pays développés ont presque toutes un système informatisé de gestion, en particulier les modules qui assurent la création du catalogue et le prêt. L'informatisation fait disparaître inéluctablement l'édition des catalogues sur fiches au profit des catalogues interrogeables en ligne, les OPAC (Online Public Access Catalogues [Catalogues en accès public en ligne]). Depuis les années 90, le développement de l'Internet et récemment de l'Intranet, tend à faire disparaître la distribution des catalogues sur CD-ROM au profit d'un accès direct par les réseaux aux fichiers constamment à jour.

### L'interconnexion des catalogues

L'adoption de règles de catalogages et de formats suffisamment compatibles permet de réaliser l'interconnexion des catalogues par la mise en œuvre du protocole commun de communication Z39.50 (ISO 23950). Ce protocole est une norme de communication entre logiciels et règle le dialogue entre l'ordinateur appelant, le poste client, et l'ordinateur appelé, le poste serveur. Ce dialogue suppose notamment que la définition des champs indexés soit la même entre les deux systèmes. Ainsi, le lecteur d'une bibliothèque particulière peut interroger n'importe où dans le monde les catalogues fonctionnant sous ce protocole, tout en utilisant l'ergonomie et la langue de dialogue qui lui sont familières, celles de sa propre institution.

### Les collections numériques

Depuis quelques années, les grandes bibliothèques se sont lancées dans des programmes de constitution de collections numériques pour des raisons de préservation ou d'élargissement de l'accès aux documents rares ou patrimoniaux. Ce mouvement est tellement important que des

sites sur les collections numériques se sont créés sur l'Internet. Le mouvement a été récemment accéléré par le programme du G7 Memoria universalis dont l'Unesco est, avec la Commission européenne, membre observateur.

(<http://portico.bl.uk/gabriel/bibliotheca-universalis/digit.htm>)

Parmi les programmes emblématiques on peut mentionner les suivants :

- L'INIST (Institut national de l'information scientifique et technique), organisme documentaire français, a été le premier à numériser systématiquement 1 500 titres de revues scientifiques internationales pour automatiser son processus de fourniture de documents. Le programme a commencé en 1990, et continue toujours, à raison de 1,5 million de pages par an. ([www.inist.fr](http://www.inist.fr))
- La Bibliothèque nationale de France a engagé un programme de numérisation de 100 000 ouvrages (30 millions de pages, pour l'essentiel en mode image) de littérature française et de 300 000 iconographies. Les documents libres de droits sont mis en accès sur l'Internet. ([www.bnf.fr](http://www.bnf.fr))
- Le National Digital Library Project a commencé en 1995 et doit s'achever en 2000. Il s'agit d'un projet coopératif qui implique de nombreuses grandes bibliothèques de recherche, piloté par la Bibliothèque du Congrès. Il se propose de numériser un million de documents patrimoniaux (textes, photographies, enregistrements sonores et films) sur les sources de l'histoire des États-Unis d'Amérique. Ces documents sont mis en accès sur l'Internet. (<http://lcweb2.loc.gov/amhome.html>)
- Archivo general de Indias est un programme de numérisation de 10 millions de documents des archives de Séville, en partenariat avec IBM. ([www.mcu.es/lab/archivos/AGI.html](http://www.mcu.es/lab/archivos/AGI.html))
- La politique de numérisation des bibliothèques allemandes. Le Deutsche Forschungsgemeinschaft (DFG) subventionne depuis 1997 des programmes de numérisation des collections de bibliothèques universitaires allemandes.

Le projet national répartit la responsabilité de numérisation par tranches chronologiques et par thème entre différentes bibliothèques exactement comme les Sonder Sammel Gebiete qui répartissent les acquisitions partagées par domaine de recherche entre les bibliothèques universitaires. En 1997, deux centres techniques de numérisation ont été créés l'un à Munich et l'autre à Göttingue ([www.gbv.de](http://www.gbv.de)), qui testent les scanners, développent des standards bibliothéconomiques et techniques, étudient les systèmes de gestion de documents numériques et produisent des recommandations à l'usage des bibliothèques allemandes. Réalisée par la Fédération internationale des associations de bibliothécaires et des bibliothèques (IFLA) au nom de l'UNESCO, une grande enquête sur les programmes de numérisation dans les bibliothèques et autres institutions culturelles est en cours. ([www.ifla.org/VI/2/p1/miscel.htm](http://www.ifla.org/VI/2/p1/miscel.htm))

L'édition électronique scientifique (voir aussi l'encadré 7.2)

Parallèlement à l'édition électronique commerciale progressivement mise en place depuis cinq ans par les grands éditeurs internationaux (Elsevier, Springer, Academic Press, Kluwer, etc.), les communautés de chercheurs et les bibliothèques universitaires développent des filières d'édition électronique pour leur propre production scientifique (preprints, thèses, colloques, etc.) et pour alimenter les technologies d'enseignement à distance (cours, didacticiels, etc.). Ces productions sont traitées dans des formats plus élaborés que le mode image qui a prévalu pour les opérations de numérisation des documents papier. Il s'agit ici de formats de type SGML (Standard Generalized Markup Language), en

particulier HTML et XML (Extensible Markup Language). La DTD (Description du type de document) de la TEI (Text Encoding Initiative) est la base de toutes les opérations récentes de structuration des documents universitaires en sciences humaines et sociales en Amérique et en Europe. Le protocole Z39.50 a désormais des extensions qui permettent d'exploiter directement la structure de type SGML pour la recherche dans le document.

L'acquisition des documents électroniques publiés

L'offre de documents électroniques pose actuellement de nombreux problèmes aux bibliothèques. Certaines publications ne paraissent que sous cette forme, il est donc évident que les bibliothèques doivent les acquérir ou au moins y accéder. Cependant la réglementation qui se met en place assimile l'acquisition du document électronique à une licence d'utilisation du contenu et fait l'objet d'un contrat bilatéral. Les questions juridiques sont actuellement portées sur le devant de la scène internationale (voir le chapitre 8). Les bibliothèques ont essentiellement recours à un accès aux documents sur le serveur de l'éditeur ou à travers un prestataire qui peut être une agence d'abonnement (Swets, Dawson, Blackwell's, etc.) ou d'un réseau coopératif (OCLC, Pica, etc.). La livraison de documents électroniques sur le site local d'une université ou d'un consortium de bibliothèques est encore expérimentale, mais semble être la solution d'avenir pour l'exploitation de la littérature scientifique éditée.

CHRISTIAN LUPOVICI,  
Directeur de la bibliothèque  
de l'Université de Marne-la-Vallée (France)

relativement mal connue. De plus, l'accès au contenu des plus anciens des supports modernes est subordonné à l'emploi de dispositifs de restitution qui sont actuellement en voie de disparition parce qu'ils sont de moins en moins demandés et que les pièces de rechange sont introuvables. La numérisation de l'écrit, de l'image et du son offre de nouvelles perspectives de préservation et de stockage mais on ignore aussi

jusqu'à quel point les supports numérisés sont durables. Il est possible, toutefois, d'assurer un accès au contenu des documents d'archives par la numérisation, ce qui évite l'usure et la possible dégradation des documents originaux. Les supports numérisés sont assez nouveaux, et aucune régulation ou norme générale concernant leur emploi à des fins de préservation n'a encore été définie.

### Outils de recherche

Dans les bibliothèques, les catalogues ont figuré parmi les premiers éléments à être informatisés. Au début, cette informatisation a facilité le processus d'impression, puis elle s'est étendue à l'établissement de catalogues en ligne. Les grands index et les recueils de résumés analytiques ont connu une évolution similaire, de l'impression traditionnelle à l'impression par ordinateur, puis à la disponibilité en ligne et sur CD-ROM. L'informatique a aussi permis le développement d'outils de recherche tels que les index de citations et les concordances qui sont produits automatiquement. Un grand nombre de processus de recherche tels que les opérateurs booléens ont été élaborés et affinés au cours des dernières années. Ils sont très coûteux, de même que le temps de connexion pour les bases de données. Les prix en question sont prohibitifs pour de nombreuses bibliothèques, et la conclusion de contrats entre les vendeurs et toutes les bibliothèques d'un pays ou de grands groupes d'établissements non commerciaux est une des façons de résoudre ce problème. Ces types de négociation et de contrat, qui couvrent aussi les vendeurs de revues électroniques et d'autres produits, deviendront de plus en plus importants à l'avenir pour les établissements qui offrent des services d'information.

### Utilisation de documents en texte intégral, de son et d'images

La capacité de stockage de données des ordinateurs permet de stocker non seulement les indications bibliographiques relatives à une publication, mais aussi le texte complet de celle-ci. D'anciens textes, initialement imprimés ou même manuscrits, sont scannés et numérisés de sorte que les œuvres de grands écrivains sont mises à la disposition de tous. Les textes peuvent être lus page par page, ou certaines parties d'entre eux peuvent être sélectionnées grâce à une recherche fondée sur une combinaison de mots.

En ce qui concerne les documents en texte intégral, les enregistrements sonores et les images, les questions de droit d'auteur figurent, en raison de leur complexité, parmi les principaux problèmes auxquels les établissements fournissant des services d'information ont à faire face. Dans la plupart des pays, la réglementation du droit d'auteur s'est développée au long de plusieurs siècles, et l'utilisation de documents imprimés par les bibliothèques est réglementée d'une façon qui convient à toutes les parties. Mais rien de tel n'a encore été mis en pratique pour les supports numériques, à la fois parce qu'ils existent depuis peu et parce qu'il est difficile de les protéger contre la copie illicite. Les enjeux financiers sont énormes, et les copier est très facile et assez peu coûteux pour les particuliers, la technique étant relativement simple. Les faits nouveaux intervenus récemment dans le domaine du droit d'auteur et des droits voisins sont passés en revue au chapitre 8.

### Revue électronique

Vingt-sept revues électroniques, dont sept faisaient l'objet d'examen par les pairs, figuraient dans la première édition du *Répertoire des revues, bulletins d'information et listes de discussion entre spécialistes*, qui date de 1991. L'édition de 1997 énumère 2 500 revues, dont plus de 1 000 font l'objet d'un examen par les pairs (voir l'encadré 7.2).

Des services de revues électroniques ont été offerts en 1997 par de grandes agences. Swets et Zeitlinger a déjà été mentionné comme l'agent du projet du Royaume-Uni, mais Blackwell offre aussi ce type de service, de même que le plus grand éditeur de revues scientifiques, Reed Elsevier, qui se limite toutefois aux siennes.

La base de données en texte intégral en ligne sur les collections électroniques d'OCLC (On Line Computer Library Center), également créée en 1997, est un service fondé sur le Web qui porte sur plus de 1 000 revues publiées par 30 éditeurs.

## Gestion des documents

Toutes les procédures administratives des établissements fournissant des services d'information peuvent être automatisées. Les documents comme les utilisateurs peuvent être identifiés par des codes ; les documents peuvent être commandés automatiquement et les factures relatives aux recherches en ligne sont imprimées par ordinateur. Toutes ces opérations ont une dimension archivistique, les établissements devant être en mesure de retrouver les données qui ont un intérêt historique, et ceux du secteur public étant généralement tenus de conserver leurs dossiers quelques années avant de les archiver. Il existe une certaine demande de systèmes de sauvegarde et de copie de sécurité dans le cas du traitement électronique. Même le choix des livres et des autres composants des fonds détenus par les bibliothèques se fait électroniquement. Les responsables des archives doivent prendre des décisions judicieuses au sujet de la présentation des données fournies, des supports de stockage, des principes de sélection, etc.

## L'INTERNET

Les modes de travail dans les bibliothèques et les archives sont transformés par deux révolutions, la première avec l'arrivée des ordinateurs et la seconde avec celle de l'Internet. Il est maintenant possible de rechercher des informations sans se soucier des fuseaux horaires et de l'emplacement géographique où elles sont stockées et d'obtenir les données les plus à jour parce que les processus d'impression et d'édition ne sont plus nécessaires ou sont très simplifiés.

Le courrier électronique est un moyen de communication qui remplace dans une large mesure le courrier traditionnel et la télécopie employés auparavant par les établissements tout en prenant moins de temps à l'utilisateur. Le Web — on dit aussi la Toile — peut être considéré comme une gigantesque base de références, où tous les types d'information sont disponibles sous forme de métadonnées ou de texte inté-

gral, de sons et d'images. Les questions liées à l'utilisation illicite des technologies de l'information, et plus particulièrement de l'Internet, sont examinées au chapitre 8.

Le Web est mis à profit pour de nombreuses activités quotidiennes impliquant une transmission d'information qui, auparavant, ne pouvaient être obtenues qu'en consultant des textes imprimés ou des télétextes et, comme le montre le chapitre 2, on trouve sur l'Internet de nombreux types de sources utiles d'information pour l'enseignement et la recherche. On peut dire que ces activités ont établi les bases de réseaux mondiaux d'information et de communication électronique. Chaque établissement a également la possibilité de créer son propre site Web pour informer les utilisateurs des services qu'il offre et leur faire connaître son mode d'organisation. Différents sites Web sont reliés de sorte qu'un point d'entrée ou d'accès conduit à plusieurs autres sites.

L'Internet permet aussi d'accéder aux bibliothèques et à leurs catalogues. Le plus récent répertoire de catalogues interrogeables en ligne, qui date de 1998, est un guide des catalogues figurant sur l'Internet. Il contient 1 434 notices en provenance de l'ensemble du monde, mais dont 888 ont pour origine les États-Unis d'Amérique et la plupart des autres l'Australie, le Canada, la Nouvelle-Zélande et l'Europe de l'Ouest. Les pays en développement et ceux d'Europe de l'Est ont peu de notices ou n'en ont aucune. Les informations données par le guide ont été rassemblées au printemps de 1996 et, depuis lors, le nombre des bibliothèques ayant de tels catalogues accessibles par l'Internet s'est beaucoup accru en Europe de l'Ouest et au Canada. Le Danemark, par exemple, avait onze notices concernant toutes des bibliothèques universitaires dans le guide de 1998. A la fin de cette année-là, plus de 125 bibliothèques danoises, dont 75 étaient publiques, avaient de tels catalogues.

L'Internet a rendu possibles de nouvelles activités. Des conférences électroniques ou des groupes informels de discussion peuvent être organisés avec un

## Encadré 7.2 → L'édition électronique scientifique

Les revues électroniques confèrent à la publication une valeur ajoutée qui intéresse fortement les scientifiques et les éditeurs, et le nombre de revues électroniques spécialisées dans les sciences, l'ingénierie et la médecine — qu'elles y soient ou non des références — a augmenté de manière spectaculaire ces dernières années. Cependant, l'édition électronique, avec sa plus grande souplesse et la variété de ses modes de présentation, défie les normes et les pratiques habituelles. Comment la culture et les pratiques en cours dans le monde des publications scientifiques vont-elles en être affectées ? Quelles normes faut-il appliquer dans l'environnement électronique ?

C'est afin d'explorer ces questions parmi d'autres qu'a été organisé, du 12 au 14 octobre 1998, sous les auspices de l'American Association for the Advancement of Science (AAAS), du Conseil international des unions scientifiques (CIUS) et de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO), un atelier international qui avait pour objet d'examiner l'application des méthodes électroniques à la publication des revues scientifiques en vue de faciliter l'élaboration de pratiques et de normes reconnues au niveau international. L'atelier devait s'appuyer sur les travaux de la conférence internationale d'experts sur « l'édition électronique en sciences » qui avait été réunie par Les Presses du CIUS et l'UNESCO en février 1996. Une recommandation majeure de cette réunion réclamait la convocation d'urgence d'un forum rassemblant des scientifiques et leurs organisations « pour formuler des codes d'éthique et de conduite applicables à la publication électronique qui énonceraient les obligations réciproques du scientifique et de la collectivité sur des questions comme l'examen critique par les pairs, la citation, l'intégrité et l'authentification du matériel et l'archivage ».

La grande diversité des pratiques et des traditions qui ont cours dans le domaine des publications savantes de toutes les disciplines a de quoi inciter à la prudence quiconque cherche à élaborer des principes généraux destinés à recevoir un vaste champ d'application. Néanmoins, l'atelier a pu identifier un certain nombre de problématiques qui méritent de retenir l'attention de la communauté scientifique dans son ensemble. Ce sont les suivantes.

**Définition d'une publication.** Étant donné que le traitement numérique facilite la production et la conservation de plusieurs versions publiques d'un document ou d'un article scientifique, l'atelier a recommandé que chaque version d'un document mise à la disposition du public porte, bien en vue, la

mention explicite de son état, formulée de manière directement compréhensible.

**La citation.** En raison de l'existence possible de multiples versions d'un document, l'atelier a recommandé que la communauté scientifique s'emploie à mettre au point, en matière de citation, des pratiques normalisées qui soient adaptées à la science, comprennent des métadonnées appropriées, soient aptes à une affectation automatique et faciles à utiliser.

**L'appréciation critique des pairs.** La facilité de publication qui caractérise les médias électroniques renforce la nécessité de mesures adéquates de contrôle de la qualité en ce qui concerne tant le contenu scientifique que la présentation. Les sociétés et les revues scientifiques devraient par conséquent établir et diffuser des directives visant à maintenir la qualité et l'intégrité du processus d'examen critique.

**La fraude scientifique.** Toute réduction des mesures de contrôle applicables à la publication électronique accroît les risques de fraude scientifique, et notamment de falsification des résultats ou de plagiat, bien que les caractéristiques techniques de la publication électronique (comme la facilité de balayage automatique et de recherche des similitudes) augmentent les chances de détection des fraudes. Il faut donc encourager la recherche sur l'application des méthodes électroniques à la détection de la fraude scientifique.

**Le libre accès.** L'atelier a étudié le conflit entre le besoin qu'ont les scientifiques d'accéder directement à de grandes bases de données et à des collections d'observations scientifiques, et la nécessité pour les compilateurs de voir leurs efforts recevoir une rétribution commerciale. Les résultats des recherches financées par des fonds publics devraient être clairement reconnus comme « biens publics », et l'accès libre et complet aux données rassemblées est jugé indispensable au progrès scientifique. Il ne s'ensuit pas nécessairement que la « gratuité » de l'accès doive être décrétée au bénéfice des chercheurs, car le rôle que jouent les éditeurs, les compilateurs, les bibliothécaires et autres facilitateurs doit recevoir sa juste récompense pour leur permettre de continuer à remplir utilement leurs fonctions (voir aussi le chapitre 8). L'atelier a appelé l'attention de la communauté scientifique, des organismes de financement et des législateurs sur le fait que l'entreprise scientifique est vitalemment tributaire de la capacité des chercheurs à faire usage de collections de faits et d'observations, et que les mesures qui limitent l'accès au

matériel utile sont contraires au bien public. Les législateurs sont incités à mettre au point un mécanisme permettant de faire un « juste usage » des grandes bases de données afin de promouvoir l'accès libre et complet à des données essentielles pour la recherche et l'enseignement scientifiques, moyennant des incidences négatives minimales sur les intérêts commerciaux des détenteurs des droits.

Le domaine privé. Il est maintenant possible pour les rédacteurs et les éditeurs de revues de rassembler des informations détaillées sur la nature du matériel accessible et son usage, et de compiler des profils d'auteurs/utilisateurs. Les types d'informations rassemblées ainsi que l'usage qui en est fait devraient être déclarés en toute netteté par les rédacteurs en chef de revues et/ou les éditeurs. Il ne saurait être permis de divulguer à qui que ce soit des informations précises concernant les personnes sans l'autorisation des intéressés.

L'archivage. Un frein très important à l'adoption des méthodes électroniques de publication est le manque de

moyens d'archivage car, en l'état actuel des progrès de la technologie, il n'est guère possible de garantir aux auteurs et aux éditeurs que le matériel publié par voie électronique restera disponible et lisible au cours des décennies et des siècles à venir. La création d'archives pour la conservation de longue durée des publications électroniques devrait être encouragée à titre de facteur fondamental du développement scientifique et culturel.

Les pays en développement. L'édition électronique offre aux pays en développement une occasion unique de faire avancer chez eux la communication scientifique. Il est possible grâce à elle d'améliorer l'accès à la littérature mondiale, de combler les manques des collections locales et d'améliorer la visibilité de leurs propres contributions scientifiques.

Source : AAAS/UNESCO/CIUS. Atelier sur la définition de pratiques et de normes pour l'édition électronique dans le domaine des sciences. Rapport final (en anglais seulement). [www.aaas.org/spp/dspp/sfr1/projects/epub/report.htm](http://www.aaas.org/spp/dspp/sfr1/projects/epub/report.htm)

accès limité ou ouvert, et surfer sur le Web est devenu l'équivalent de zapper à la télévision ou de feuilleter des livres sur les rayons d'une bibliothèque.

L'Internet faisant appel à des techniques sensibles et vulnérables, les possibilités de dysfonctionnement sont nombreuses. Les établissements qui desservent un grand nombre d'usagers ou qui ont entrepris des programmes ambitieux et coûteux tels que la numérisation massive risquent tout particulièrement d'en souffrir, mais cela peut aussi avoir des conséquences catastrophiques pour les petits établissements. Le trafic sur le Web et le nombre de sites Web connaissent une croissance exponentielle, ce qui soulève des problèmes de files d'attente et de délais de réponse inacceptables.

Pour les professions de bibliothécaire et d'archiviste, d'autres aspects du Web posent problème. La question centrale est, pour tous les documents électroniques, celle de leur authenticité (voir aussi le chapitre 11). Peut-on jamais être certain que le document a effectivement été produit par la personne ou l'organisation qui affirme en être l'auteur ? Le texte est-il aujourd'hui le même qu'hier et sera-t-il encore le même demain ? Une thèse comportant des renvois à des documents électroniques non imprimés est expo-

sée au risque qu'ils soient modifiés ultérieurement ou même disparaissent. La question est étudiée par des spécialistes qui élaborent des normes pour ces documents. L'authenticité est aussi un problème pour les services d'archives. La grande quantité de données que l'on peut avoir à explorer représente une autre difficulté. Une recherche fondée sur un ou quelques mots clés débouche aisément sur plusieurs milliers d'adresses. La recherche peut être affinée en utilisant les outils élaborés à cette fin, les opérateurs booléens, les limitations linguistiques ou géographiques, la provenance, etc. Mais le Web étant, de ce point de vue, complètement anarchique, cela ne résout pas tous les problèmes. Une bonne partie des textes qu'on y trouve n'ont pas fait l'objet d'une mise au point rédactionnelle et aucune règle générale n'a été définie dans des domaines tels que la terminologie. Les explorations à des fins d'enseignement et de recherche peuvent fournir un grand nombre d'exposés subjectifs dont la rigueur n'est assurée par aucun contrôle de qualité. Les textes imprimés ou publiés sur d'autres supports ont généralement été soumis à un processus de sélection fondé sur des critères universitaires, professionnels ou économiques, mais il n'en va pas de même pour les produits d'« information » artisanaux.

Les spécialistes de l'information ont beaucoup à faire pour mettre en œuvre le concept de sélectivité. Cela pose toutefois le problème de la mesure dans laquelle la censure est préférable à la confusion et à l'inutilité. Certaines de ces questions sont examinées dans une perspective plus générale au chapitre 11, qui est consacré à l'Internet.

## LA BIBLIOTHÈQUE VIRTUELLE

L'idée de rassembler toutes les connaissances humaines est très ancienne, et on la retrouve dans de nombreuses civilisations, comme en témoignent la Bibliothèque de l'ancienne Alexandrie et les grandes entreprises encyclopédiques des civilisations arabe, chinoise et européenne. Le but est de parvenir à une vision d'ensemble de tout le savoir fixé sur un support grâce, par exemple, au système de la Classification décimale universelle ou au moyen d'un grand ordinateur. Récemment, et du fait de la croissance de l'Internet, cette vision s'est à nouveau manifestée sous la forme de « la bibliothèque sans murs », la bibliothèque virtuelle.

L'hypothèse est la suivante : tout le monde peut avoir instantanément accès à la totalité de l'information fixée sur un support où que ce soit dans le monde grâce à la bibliothèque la plus proche. Ce phénomène a été décrit de plusieurs manières. On parle assez souvent de bibliothèque électronique, mais il pourrait aussi bien s'agir d'une grande base de données ou d'un nouveau développement de l'automatisation des bibliothèques, ou encore d'applications de l'hypertexte sur le Web. L'expression que la profession préfère est bibliothèque numérique, ce qui n'est pas seulement le nom d'une collection de supports numérisés. Il s'agit avant tout d'une **bibliothèque**, qui ne se situe pas nécessairement dans tel ou tel bâtiment, mais est une **organisation systématique**, où diverses opérations professionnelles sont exécutées par des spécialistes à l'intention de groupes précis d'utilisateurs. La collection, ou des parties de celle-ci, peut avoir été numérisée, et l'accès à des réseaux électroniques ou autres

assuré. Les **métadonnées**, qui décrivent le contenu et les attributs des collections, sont un des éléments les plus importants des bibliothèques numériques. Elles facilitent en effet l'exploration de millions de documents. L'exemple le plus marquant à cet égard est ce qu'on appelle le **Dublin core** (noyau de Dublin), qui définit quinze éléments centraux à utiliser par les auteurs comme par les intermédiaires. Un autre problème est celui de la désignation ou, autrement dit, des chaînes de caractères qui identifient de manière univoque chaque objet numérique et font donc partie des métadonnées des documents. Ce système de désignation doit être permanent, ce qui signifie qu'un nom ne peut être lié à un emplacement déterminé. Sa création n'est pas une question technique, mais un problème d'organisation, et différents systèmes ont été suggérés, comme les langages répondant aux besoins constants des utilisateurs (Persistent User Requirement Languages [PURL]) élaborés par l'OCLC, où un serveur recherche l'emplacement uniforme de ressource (Uniform Resource Locator [URL]) dans une base de données, le nom uniforme de ressource (Uniform Resource Name [URN]) élaborée par le groupe de travail technique de l'Internet (Internet Engineering Task Force [IETF]), qui n'est pas encore exploitée, et le système de l'identificateur d'objets numériques (Digital Object Identifier [DOI]), mis au point par l'Association des éditeurs américains et la Corporation for National Research Initiatives. Ces systèmes fournissent un moyen d'identifier les objets numériques et d'y avoir accès, et leur élaboration a été stimulée par les enjeux de la gestion du droit de propriété intellectuelle.

Les bibliographies nationales sont les pierres angulaires du système de bibliothèques et leur élaboration se heurte à de nombreuses difficultés inhérentes aux documents électroniques. Certains de ces documents, comme ceux dont les adresses comportent, par exemple, « com », « org » et « net », posent les mêmes problèmes que les textes imprimés, mais d'autres complications découlent de l'existence même de documents dynamiques. Un projet pilote entrepris récem-

ment en Suède vise à tester des méthodes de collecte et de préservation de documents électroniques en ligne et de fourniture d'accès à ces documents de manière à leur permettre d'être considérés comme publiés. A cette fin, on prend un « instantané » des pages HTML (Hypertext Markup Language) pertinentes, mais cela ne concerne que les documents électroniques statiques. L'Internet n'est pas la bibliothèque virtuelle mais en constitue une partie importante. Depuis quelques années, les bibliothécaires parlent souvent du passage « de la collection à la connexion ». C'est sans nul doute une description très précise de l'évolution intervenue, mais la vision de la bibliothèque virtuelle ou numérique doit englober la « collection » au même titre que la « connexion ».

## PERSONNEL ET ENSEIGNEMENT

Il est utile de se référer à un secteur unique des services d'information du fait des similitudes entre les missions des établissements qui s'y trouvent et entre les tâches de leur personnel. Les archives préservent pour les générations futures des documents allant de l'exemplaire unique d'un texte manuscrit ayant une grande valeur historique à des échantillons de documents produits en masse, tels que les formulaires de recensement ou les déclarations d'impôts. A cette fin, des méthodes d'échantillonnage statistique ont été élaborées pour que le nombre d'objets préservés suffise à répondre aux diverses préoccupations des chercheurs. Les distinctions entre le travail intellectuel, technique et administratif et entre la classification et l'archivage des documents ont été abolies, de sorte que le domaine de l'archivistique s'est étendu et que ses approches systématiques se sont élargies. La personne qui accomplit ce travail ainsi redéfini est généralement appelée gestionnaire de documents. Les bibliothèques, quant à elles, rassemblent des collections de textes imprimés et ont élaboré des systèmes de recherche fondés sur des catalogues et des systèmes de classification. Le travail des bibliothécaires comprend aussi

le choix des ouvrages à acheter ou à éliminer. On s'attache de plus en plus à mieux servir les utilisateurs, et des dispositifs tels que des catalogues collectifs et des mécanismes de prêt entre bibliothèques ont été mis en place à cet effet. L'utilisation des technologies de l'information et de la communication a permis de développer encore plus ces systèmes. Le troisième groupe d'établissements, les entreprises fournissant des services d'information, a ses origines dans le mouvement de développement de la documentation qui s'est constitué à partir du début du siècle. Il s'attache principalement à fournir des services aux responsables du secteur des entreprises et des affaires. Ces entreprises utilisent fondamentalement les mêmes outils que les bibliothèques mais veillent tout particulièrement à ce que les informations soient fournies rapidement, et les technologies de l'information et de la communication leur facilitent beaucoup la tâche. Tous les types de personnel de ces trois groupes d'établissements accomplissent des tâches analogues pour assurer le fonctionnement de leur institution ou entreprise.

La caractéristique commune des trois domaines, de leurs établissements et de leurs professionnels est le rôle d'intermédiaires qu'ils jouent entre les documents (imprimés ou électroniques) et les utilisateurs. Les technologies de l'information et de la communication ont brouillé les frontières entre les différents types de documents. Les ordinateurs sont utilisés à des fins administratives et les réseaux électroniques se sont étendus, de sorte que d'autres types d'établissements tels que les musées ont également des activités analogues. De ce fait, les spécialistes sont progressivement remplacés par des généralistes capables de maîtriser les technologies de l'information et de la communication ou par de nouveaux professionnels, tels que des concepteurs de pages d'accueil et des administrateurs de site dont la formation associe souvent l'esthétique aux sciences de l'information. La préparation à l'emploi dans ce secteur tend généralement, à mesure qu'elle se développe, à privilégier davantage les technologies de l'information et de la communi-

cation, et tout d'abord leur utilisation, mais plusieurs établissements d'enseignement mettent également l'accent sur les compétences en matière de programmation et le savoir-faire technique. Les ordinateurs sont utilisés dans tous les domaines de l'enseignement et de la recherche, et les étudiants se familiarisent, grâce à leurs adresses électroniques et à leurs sites Web, avec les exigences d'une utilisation régulière de la technologie et les possibilités qu'elle offre.

Les cours de bibliographie portaient essentiellement sur les bibliographies imprimées, complétées par des bases de données en ligne. Aujourd'hui, c'est l'inverse. L'enseignement professionnel insiste par ailleurs plus sur la théorie et la méthodologie que sur les compétences pratiques. En effet, l'évolution est si rapide que des changements importants interviennent entre le début et la fin des études. Les connaissances méthodologiques restent valables plus longtemps et assurent une plus grande adaptabilité du personnel.

La convivialité est un aspect important des technologies de l'information et de la communication, mais la formation des utilisateurs dans les bibliothèques reste très nécessaire. Ceux-ci ont désormais accès, grâce à l'électronique, aux ressources des bibliothèques sur place, ou depuis leur bureau ou leur domicile. Ils doivent donc mieux maîtriser les systèmes numériques, et les différents catalogues publics en ligne ne doivent plus être conçus pour les seuls professionnels. L'évolution des tâches accomplies par le personnel et des programmes d'enseignement est illustrée par les désignations et les titres utilisés. Ainsi la bibliothéconomie a cédé la place à la science des bibliothèques, puis aux sciences des bibliothèques et de l'information, et enfin à la gestion des systèmes d'information. Parmi les demandeurs d'emploi dans ce domaine figurent notamment des administrateurs et des concepteurs de sites.

Un certain nombre de nouveaux emplois de spécialistes ont été créés dans les secteurs privé et public pour des personnes capables de maîtriser les technologies de l'information et de la communication. De

plus, leur formation traditionnelle, avec son approche systématique, est un atout. Les créateurs de logiciels, les concepteurs de systèmes d'information pour tous les types d'entreprises et d'organisations, les consultants pour les questions relatives à l'Internet, les enseignants, les formateurs et le personnel des services commerciaux des établissements se consacrant à l'information ne ressemblent plus guère aux archivistes et aux bibliothécaires que l'on connaissait. La conjoncture est très favorable aux professions de l'information, mais on peut évidemment se demander si, avec la récession économique, locale ou plus générale, cet optimisme ne sera pas contredit par les faits. Il est néanmoins clair que le développement rapide des technologies en question exigera que les professionnels de l'information, comme ceux de nombreux autres domaines, continuent à étudier tout au long de leur vie et qu'établissements d'enseignement et associations professionnelles mettent en place les programmes et les moyens voulus.

## UN AVENIR PARADOXAL

Certains paradoxes sont inhérents à l'utilisation croissante de ces technologies. D'un côté, l'accès à l'information et à la communication devient plus facile, et les ordinateurs personnels sont meilleurs, plus rapides et moins onéreux. Ces facteurs peuvent être considérés comme mondialement porteurs de démocratisation. De l'autre, le développement de la technologie, des logiciels et du traitement des données exige des investissements de plus en plus importants, ce qui favorise la création à l'échelle mondiale de monopoles de la production et de la vente.

L'apparition d'un nouveau type d'analphabètes, qui ont découvert les technologies de l'information et de la communication à l'âge adulte et sont donc désavantagés sur le marché de l'emploi, est un autre paradoxe résultant des progrès accomplis en matière d'accessibilité. De ce point de vue, l'utilisation des ordinateurs personnels peut être comparée à la maîtrise d'une langue étrangère, qui exige que l'on

s'exerce régulièrement à parler, à lire et à écouter. Les bibliothèques peuvent jouer un rôle important pour lutter contre cette forme d'analphabétisme en assurant à tous les utilisateurs un accès à l'ordinateur personnel et aux réseaux. Dans certains pays d'Europe de l'Ouest, la question de savoir si les bibliothèques publiques devraient offrir à tous les utilisateurs un accès gratuit à l'Internet fait actuellement l'objet de débats.

C'est principalement dans le cadre des pays et des économies capitalistes que les technologies de l'information et de la communication se sont développées, notamment parce que la substitution de machines à une main-d'œuvre qui revient relativement cher réduit le coût de la production et des services. Mais l'évolution est si rapide qu'on manque partout de personnel qualifié. Cela donne aux professionnels compétents la possibilité d'être très bien rémunérés et de travailler dans d'excellentes conditions. Les spécialistes manqueront aussi en Europe de l'Est et dans le monde en développement à mesure que ces technologies s'y diffuseront. La situation pourrait même y être pire si, comme c'est le cas pour les artistes ou les sportifs, les rémunérations plus élevées incitent le personnel qualifié à venir en Europe de l'Ouest ou en Amérique du Nord. On peut également se demander si la prédominance de la langue anglo-américaine dans la communication internationale ne sera pas encore renforcée. Cette question est examinée au chapitre 3.

La qualité de l'information passe par le problème de la recherche de données pertinentes et exactes dans les millions de documents disponibles sur l'Internet. Il y a quelques années, les grandes bases de données étaient déjà surnommées systèmes GIGO (*garbage in-garbage out*, ce qui peut être rendu par « foutaises en entrée, foutaises en sortie »), et le problème n'a fait que s'aggraver à mesure qu'on reliait « toutes » les bases de données. C'est un des principaux défis que les professionnels de l'information de l'avenir devront relever. Les possibilités de modifier les documents électroniques se traduisent pour la profession par

d'autres difficultés. L'authenticité doit être définie pour résoudre les problèmes d'archivage, de recherche et de droit de propriété intellectuelle. Avec la numérisation, l'imprimé, l'image et le son se stockent et se transmettent sous la même forme matérielle (voir le chapitre 10). Cette convergence se retrouve dans celle des établissements et des professions puisque tout le travail qui s'y fait a trait à des supports stockés et échangés de la même façon. A l'avenir, on pourrait fort bien constater aussi une convergence entre les producteurs d'information et les intermédiaires.

## ORIENTATIONS BIBLIOGRAPHIQUES

CLEVELAND, G. 1998. *Digital libraries : definitions, issues and challenges*. IFLA UDT Core Programme, Occasional Papers 8.

FEATHER, J. 1998. *The information society. A study of continuity and change*. 2<sup>e</sup> éd. Londres, LA Publishing.

Des articles sur les nouveaux développements peuvent être trouvés dans de nombreux périodiques parmi lesquels on peut citer les suivants : *Records Management Journal*, *The Electronic Library*, *International Journal on Digital Libraries*, *Internet Research*, ainsi, bien sûr, que sur les sites Web des organisations et entreprises actives dans ce domaine.