

De la difficulté à utiliser les bibliothèques numériques

Par Alain Jacquesson*

Les NTIC sont majeures depuis longtemps

Alors que les bibliothèques numériques existent maintenant depuis près de dix ans, il reste toujours difficile d'en donner une définition admise par tous. La nature numérique des documents qui les composent est naturellement un élément essentiel. En regardant en arrière, on constate que ces nouveaux objets sont apparus de façon très diffuse. Historiquement les premiers documents numériques n'ont pas été créés et déposés dans les bibliothèques ; ils ont été conçus au moins dans trois endroits différents : chez les éditeurs qui, dès la fin des années soixante, voulaient alimenter leurs photocomposeuses avec des textes réutilisables en vue de tirages ou rééditions successives ; ensuite, des chercheurs universitaires, dès le début des années soixante-dix, ont encodé sur supports informatiques des textes classiques afin d'établir des concordances ou pratiquer différentes manipulations de textes, en particulier la linguistique statistique. Ces deux premiers groupes utilisaient généralement des moyens informatiques lourds et déléguaient toute la gestion de leur corpus à des informaticiens. Ces derniers utilisaient les gros ordinateurs de l'époque et stockaient leurs informations sur de lourdes bandes magnétiques. Enfin, depuis le milieu des années soixante, certains chercheurs en sciences exactes, notamment les physiciens, qui les premiers avaient un accès direct à des mini-ordinateurs (PDP8, Nova, etc.), s'étaient mis à rédiger leurs articles sur ces machines en utilisant des éditeurs de textes prévus à l'origine pour corriger des programmes informatiques.

Après ces développements progressifs, il y eut une formidable rupture en 1980 avec l'arrivée massive de la micro-informatique. Bien que les premières machines firent leur apparition au cours de la précédente décen-

nie (Altair en 1975, Apple I en 1977), c'est le 12 août 1981 qu'IBM lançait son *Personal Computer/PC*, créant un choc auprès des tenants de l'informatique lourde. En quelques années, tous les postes de travail du secteur tertiaire seront équipés de micro-ordinateurs. Toutes ces machines seront dotées de deux types de logiciels : les tableurs et surtout les traitements de textes. Même si, avec quelque recul, on est en droit de juger que MS/DOS s'est révélé bien peu convivial, les personnes équipées de ces nouvelles machines allaient apprendre, soit de façon totalement autonome, en lisant quelques manuels, soit grâce à une formation de quelques jours, l'utilisation d'un système de traitement de texte.

Dès 1981, créer des documents numériques devenait à la portée de chacun : de l'écolier au savant.

Une autre rupture arrivait en juin 1985 : les CD-Rom. Cette année là, la firme « *TLC - The Library Corp* » présentait au congrès de l'American Library Association le fichier bibliographique de la Bibliothèque du Congrès sous le nom de « *BiblioFile* », la première application commerciale d'un CD-Rom tous domaines confondus. À l'étonnement de tous, le fichier, même partiel, de cette grande bibliothèque tenait sur un disque de plastique de 12 centimètres de diamètre. Les CD-Rom marquaient un autre progrès. L'informatique utilisait une technologie largement diffusée dans le grand public : les lecteurs de disques compacts sonores, ce qui permettait une certaine maîtrise des coûts. Ces nouveaux papyrus, comme on les a appelés à l'époque, allaient connaître le développement que l'on sait. Comme toujours, il y aura des évolutions successives : des échecs avec le *CD-I* ou des réussites avec le *DVD*. Les CD-Roms sont aujourd'hui présents aussi bien dans les grands magasins que dans les laboratoires de recherche fondamentale. Ces documents numériques sont désormais produits soit à l'échelle industrielle, soit comme documents à diffusion restreinte limitée à un cercle de spécialistes. Les graveurs de CDR permettent même une production personnalisée de document sur disque numérique, un rêve inaccessible pour l'industrie du livre imprimé.

* Bibliothèque publique et universitaire de la Ville de Genève (Suisse)

Une autre évolution technologique fondamentale est étroitement liée au développement des télécommunications, notamment des réseaux d'ordinateurs. L'histoire est désormais bien connue : un réseau informatique militaire américain, *Arpanet*, fonctionne pour la première fois en octobre 1969 sur la base des travaux de John Licklider, ainsi que ceux de Leonard Roberts. Ce réseau évolue rapidement vers un réseau reliant des ressources informatiques d'universités américaines (1970), prend le nom d'*Internet*, se développe dans les milieux académiques internationaux (1980), puis s'ouvre enfin aux particuliers et à l'économie¹.

Le dernier développement est de nature qualitative. En 1989 est conçu à Genève, au CERN (Centre européen de recherche nucléaire), le *World Wide Web*. C'est un réseau de sites informatiques qui peuvent être recherchés et atteints au moyen d'un protocole de transfert appelé *http* (*Hypertext Transfer Protocol*). Il a été créé pour permettre aux physiciens de communiquer rapidement et simplement à l'échelle planétaire. La distribution gratuite en 1993 d'un premier navigateur appelé *Mosaic* assura une partie du succès du Web. Il sera rapidement remplacé par *Navigator* de Netscape ou *Explorer* de Microsoft. Ces différentes interfaces de navigation assureront le succès d'Internet dans le grand public. C'est sur elles que bon nombre de bibliothèques numériques vont être construites.

Si nous nous sommes permis de reconstituer cette brève chronologie, c'est que nous voulions surtout montrer que les nouvelles technologies de l'information et de la communication – nos fameuses NTIC, que prônent aujourd'hui nos hommes politiques et nos économistes – sont connues depuis de longues années, même si les marchés financiers ne les découvrent qu'aujourd'hui.

Les composantes essentielles des bibliothèques numériques sont disponibles

On a donc assisté progressivement à la mise en place d'outils de nature totalement différente qui seront les composantes principales des bibliothèques numériques.

En ce qui concerne la terminologie, en quelques années, les noms les plus divers vont être utilisés pour couvrir des réalités très différentes les unes des autres et sur lesquelles personne ne s'était encore accordé : bibliothèque virtuelle, bibliothèque sans murs, bibliothèque électronique, etc.

En terme de réalisations concrètes, il existe également des approches très différentes du concept de bibliothèque numérique : pour les physiciens, d'immenses réservoirs de prétrages bruts, non organisés, non validés, non indexés, comme celui de Ginsparg à la bibliothèque du Centre de recherche nucléaire de Los Alamos, constituent la bibliothèque numérique. Pour les chercheurs travaillant dans les domaines médicaux, la bibliothèque numérique se résume souvent à l'accès aux périodiques électroniques des publications hautement contrôlées. Pour les spécialistes en sciences humaines, la bibliothèque numérique sera essentiellement composée de textes classiques ayant fait l'objet d'innombrables rééditions (*La Chanson de Roland*, *Voltaire*, etc.) qu'ils veulent pouvoir manipuler, annoter, comparer dans le cadre de leurs travaux. En lecture publique, ce sera, pour l'instant, souvent une collection de CD-Roms ou de documents anciens numérisés comme à la Bibliothèque municipale de Lisieux, encore que cette réalisation de grande qualité soit plutôt une exception. Nous pourrions citer de nombreux autres exemples que les auteurs appellent « bibliothèque numérique ».

Comme on le voit, une définition simple de la bibliothèque numérique n'est pas facile à donner.

Mais avant tout, les bibliothèques numériques sont des organisations qui offrent des services, qui distribuent l'information, qui garantissent l'accès à long terme aux documents qu'elles conservent et surtout qui disposent d'un personnel réalisant ces tâches techniques, en orientant le public vers ces nouvelles ressources.

Compte tenu du foisonnement de nouveaux moyens techniques, d'outils toujours plus performants, et surtout de la publicité que l'on fait autour, on peut facilement penser que les bibliothèques numériques sont faciles à utiliser. En écoutant certains discours, elles devraient même être plus simples à utiliser que les bibliothèques traditionnelles.

Difficultés des catalogues

Revenons quelques instants sur l'utilisation des systèmes d'accès public dans nos bibliothèques. Les OPAC comptent aujourd'hui près de 20 ans d'âge. Pourtant, ils restent toujours difficiles à utiliser.

Rappelons-nous la polémique survenue il y a maintenant deux ans sur Biblio-fr. Un scientifique de grande renommée voulait prendre l'avion pour aller dans le Michigan consulter un ouvrage français de mathéma-

tiques datant de 1971 qu'il ne trouvait pas dans les bibliothèques françaises. Tollé de la profession : l'ouvrage était dans de nombreuses bibliothèques de Paris et de Navarre, dont les catalogues étaient accessibles tant sur le Minitel que sur Internet.

Face à cette indignation généralisée, une voie discordante venait de l'Université de Bretagne-Sud où le professeur en sciences de l'information Robert Nadot, démontrait en quelques interrogations que la localisation d'un ouvrage, pour un non-professionnel, était de toute évidence très compliquée : « Le scandale n'est-il pas que, alors qu'un ouvrage est manifestement présent dans le fonds de nombreuses bibliothèques françaises, il soit aussi difficile à localiser, ou, à tout le moins, que les procédures de localisations soient aussi tordues et qu'elles ne donnent pas le même résultat. [...] Il faut être un rien masochiste pour se satisfaire des outils français, on ne peut guère que leur concéder d'exister mal. [...] ». Il terminait en comparant les systèmes disponibles à des « usines à gaz ».

La critique était sévère, mais parfaitement justifiée, car étayée par de nombreux exemples. Aux Etats-Unis, Christine L. Borgman nous disait en 1986 déjà que les catalogues en ligne étaient difficiles à utiliser ; dix ans plus tard, elle confirmait que c'était toujours le cas ! Elle démontrait notamment que les règles définies par Charles Ammi Cutter en 1904 pour l'établissement d'un catalogue-dictionnaire n'étaient pas appliquées cent ans plus tard par les systèmes d'accès public disponibles sur le marché.

Si nous n'arrivons pas, pour un problème en principe simple - des catalogues collectifs -, à construire des outils efficaces et surtout facilement maîtrisables par nos utilisateurs, que va-t-il en être pour les bibliothèques numériques ?

Catalogue des difficultés

Les bibliothèques numériques sont donc aujourd'hui nombreuses et variées. Elles sont riches de millions de documents primaires numériques de tout genre. Mais en a-t-on testé l'efficacité sur nos utilisateurs ? Le cas cité au paragraphe ci-dessus nous fait craindre le pire, car de nombreuses difficultés vont être rencontrées par le lecteur qui pour la première fois ira se promener dans nos nouvelles bibliothèques numériques.

La communication « homme-machine » est certainement un domaine beaucoup plus complexe que l'on

veut bien l'imaginer. Des chercheurs (informaticiens, psychologues, parfois même bibliothécaires) se sont penchés sur le problème. Certains grands principes ont pu être définis :

- Il faut tout d'abord identifier la diversité des lecteurs (enfants, étudiants, chercheurs, grand public, etc.) qui va interroger nos bibliothèques digitales : on ne présentera pas un texte électronique de Voltaire de la même façon à un spécialiste du dix-huitième siècle qu'à un lycéen ; dans le monde de l'imprimé, le premier ira chercher ses sources dans l'édition commentée de la « Voltaire Foundation » alors que le second prendra une édition de poche qu'il pourra annoter. Comment sera préservée cette variété des sources dans le monde des bibliothèques numériques, alors que, dans le monde de l'imprimé, le nom même d'un éditeur est une classification en soi ? Une nouvelle tradition éditoriale doit prendre racine.

- Le second principe vise à établir des règles quant à l'interface elle-même (consistance, raccourcis, retour, traitement des erreurs, etc.). Nous disposons aujourd'hui d'une interface par éditeur (Springer, Elsevier, etc.). Comme nous savons qu'il se crée entre 6 et 12 nouvelles revues électroniques par jour, nous sommes inquiets pour nos lecteurs. Ou alors nous devons passer par un fournisseur qui, comme OCLC, reformate les différentes revues pour les présenter au public sous une interface unique. Là également, le choix n'est pas simple puisque la qualité du dialogue « homme-machine » sera mise dans la balance et comparée aux aspects économiques.

- Le traitement des erreurs, dues à l'homme comme à la machine, est le troisième principe qui doit être respecté dans l'élaboration des systèmes de dialogue des bibliothèques numériques. L'affichage d'un message brut comme « Error 404² » si souvent rencontré est naturellement totalement inacceptable pour tout utilisateur d'une bibliothèque numérique quel qu'il soit.

Une autre grande difficulté des bibliothèques numériques est liée au catalogue des ressources disponibles dans leurs fonds. Va-t-on utiliser les catalogues traditionnels pour les répertorier, alors qu'on en connaît les limites comme nous l'avons vu plus haut ? Préférera-t-on utiliser des nouveaux types de catalogues comme celui du projet CORC chez OCLC ? Un outil efficace mais qui s'adresse certainement d'abord à des professionnels. Laissera-t-on nos lecteurs seuls face aux moteurs de recherche, alors qu'on sait que tous ensemble ils ne couvrent que 42 % du Web (le meilleur ne couvre que 16 % du Web) ? De plus ils ne répertorient pas ce qui se trouve derrière les barrières économiques (sites payants contrôlés par des mots de passe) des éditeurs commerciaux. Enfin, des connaissances informatiques solides seront

nécessaires à nos lecteurs. Qui, même parmi les bibliothécaires, sait dans quelles circonstances il faut choisir un format PDF, Ascii, SGML, XML, etc. pour s'approprier des textes numériques et surtout quels enjeux ces choix représentent ?

« De la psychologie des objets quotidiens »

En 1988, Donald A. Norman publiait un ouvrage sur l'utilisation des objets quotidiens. Il constatait que ceux-ci prenaient un aspect toujours plus technique et complexe ; par conséquent, on utilise une partie toujours plus limitée de ces objets quotidiens. Ce professeur au MIT cite de nombreux exemples : il est notoirement connu que le « zapping » sur une télévision est parfaitement maîtrisé par toutes les couches d'âge et toutes les strates sociales de la population quel que soit leur niveau d'éducation, mais que la programmation sur un magnétoscope de l'enregistrement d'une émission en différé est diaboliquement complexe et que cette fonction n'est utilisée que par une frange restreinte de cette même population. Le problème est analogue avec, par exemple, la programmation des renvois sur les téléphones, la mise en attente d'un appel, une conversation à trois, etc., bien que les instructions nécessaires à ces opérations soient expliquées en quelques lignes dans les annuaires téléphoniques. L'utilisation d'objets quotidiens un tant soit peu techniques est donc loin d'être simple. Une bonne partie des propriétaires de ces appareils n'utilisent qu'un sous-ensemble restreint de leurs vastes possibilités. Norman a poursuivi ses études en analysant la complexité à utiliser des micro-ordinateurs.

Sur la base de ces différentes réflexions et par analogie, Christine L. Borgman nous rend attentifs aux difficultés de tous genres que rencontrent les utilisateurs des bibliothèques numériques. Il est impératif de créer des fonds documentaires numériques qui soient d'abord simples à utiliser, avant de rechercher les performances ou une surabondance de fonctionnalités à tout prix.

À l'occasion d'un congrès tenu en 1998, les responsables d'un projet de bibliothèque numérique à l'Université d'Illinois ont présenté des résultats intéressants. Cette université offre à ses étudiants un accès gratuit, mais avec inscription préalable impérative, à une vaste bibliothèque numérique. Ces chercheurs ont cependant constaté que 83 % des utilisateurs qui s'étaient branchés sur ce serveur renonçaient dès la page d'accueil : erreur d'aiguillage ou difficulté d'accès ? Puis, parmi les 17 % restants qui poursuivaient la découverte

de ce site, une personne sur deux renonçait devant le formulaire d'inscription ! Cela signifiait que moins de 9 % des personnes qui ont eu accès à cette bibliothèque numérique ont consulté des documents primaires numériques. Ces résultats ne manquèrent pas d'inquiéter au plus haut point les responsables du projet qui d'une part reprirent l'architecture de leur site et d'autre part lancèrent une vaste campagne d'information. Les résultats passèrent alors à 18 %, ce qui naturellement reste faible et qui signifie aussi que 82 % des lecteurs ont renoncé aux documents primaires ; ils n'ont même pas passé quelques instants à sélectionner des documents pertinents. Ils sont restés bloqués à la manipulation de l'instrument.

Différentes analyses ont été faites de ces premiers résultats, mais de toute évidence une première conclusion a pu être tirée de cette expérience : quelle que soit la qualité de l'outil mis à disposition, son utilisation est complexe, même pour des étudiants de niveau universitaire et aucun projet de bibliothèque numérique ne doit être bâti sur une utilisation intuitive. Un outil de recherche documentaire peut être évalué en terme de richesse de fonctionnalités, mais il doit surtout être jugé en terme de facilité d'usage.

Il nous semble donc nécessaire d'investir non seulement dans les instruments de consultation des bibliothèques numériques mais aussi dans la formation à leur utilisation. L'emploi des bibliothèques classiques est basé sur une tradition séculaire et surtout stable. Les bibliothèques numériques sont d'abord plus complexe, puis en évolution permanente. Ces deux facteurs nous conduisent à les présenter de façon différente. Dans un article à paraître, nous avons tenté d'expliquer pourquoi nous pensons que la formation des utilisateurs sera un des défis essentiels des bibliothèques numériques.

Géographie des savoirs

De par leur définition même, les bibliothèques numériques sont dématérialisées et décentralisées. Ce qui était la propriété visible de l'institution traditionnelle – les collections matérielles – est désormais remplacé par des canaux d'accès à des informations mouvantes situées dans des territoires totalement inconnus des lecteurs (serveurs commerciaux de tout type, réservoirs numériques universitaires, etc.).

Le concept qui sera certainement le plus difficile à assimiler par nos utilisateurs concernera donc l'**architecture de l'information**. Cette notion n'est pas uni-

forme et recouvre, comme nous l'avons vu, des concepts très différents. Le premier porte sur les supports (CD-Roms, documents en ligne, etc.), le second sur les localisations : est-ce un document en ligne local (serveur du laboratoire, de l'entreprise, de la bibliothèque, etc.), accessible à distance, nécessitant une identification préalable, un abonnement, un paiement après déchargement ? Les données sont-elles homogènes (texte ou images fixes, animées, son), disponibles selon quels formats ? Des données brutes sont-elles associées aux documents ? Comment rendre homogènes et « lisibles » les rayons virtuels de nos prochaines bibliothèques ?

Ne focalisons pas notre intérêt pour les bibliothèques numériques sur les instruments d'interrogation et d'accès. Au moment de faire sa recherche, puis sa consultation, nous devons impérativement nous rappeler que le lecteur est d'abord préoccupé par son objectif et non par les outils documentaires et les chemins potentiels d'accès aux documents, car, comme nous venons de le voir, il n'y a naturellement pas une voie unique pour arriver à ces informations. Celles-ci sont dupliquées, fragmentées, recomposées, volatiles, etc. Les bibliothèques numériques et leurs nouveaux outils, certes puissants, mais incontestablement compliqués, apparaîtront alors comme autant d'obstacles à l'utilisateur. Toutes ces entraves ont souvent été créées par des informaticiens et des bibliothécaires qui focalisaient leurs préoccupations sur la puissance des instruments qu'ils créaient, mais qui étaient inconscients de la complexité de la géographie du savoir qui résultait de leurs travaux et des nouvelles contraintes économiques.

La facilité d'utilisation des bibliothèques est pour l'instant encore un mirage. Malgré les efforts réels visant à les rendre plus ergonomiques et conviviales, les bibliothèques numériques sont, et resteront longtemps encore, difficiles à utiliser.

Faut-il alors les condamner ? Certainement pas, car, de toute évidence, elles ne peuvent que croître et prendre une place toujours plus importante dans le processus de stockage et de restitution de l'information. Mais, passée la période d'euphorie qui a suivi la mise en service des innombrables bibliothèques digitales, il nous faut maintenant devenir vigilants et surtout critiques afin de ne pas accepter n'importe quoi de la part de nos fournisseurs commerciaux, scientifiques ou institutionnels.

Enfin, le rôle des bibliothécaires sera également prépondérant dans les programmes de formation des lecteurs à l'utilisation des bibliothèques numériques. Comme nous l'avons vu plus haut, la géographie des savoirs est aujourd'hui complexe et peu naturelle. Le rôle de guide, que nous devons alors assumer, sera un défi passionnant.

1. En 1971 en France Louis Pouzin fait fonctionner à l'Iria le réseau Cyclades ; certains de ces développements, notamment ceux qui concernent les protocoles, seront repris par Internet, mais Cyclades, comme d'autres réseaux européens, seront condamnés par les administrations chargées des télécommunications qui ne voyaient aucun avenir dans la communication par paquets condamnant ainsi l'Europe à un lourd retard.
2. Des pages Web qui ont été supprimées, qui ont migré vers un autre serveur, un serveur non disponible, etc.

Références :

CHRISTINE L. BORGMAN
Why are online catalogs hard to use ?
 In : *Journal of the American Society for Information Sciences*,
 37 (1986), 6 : 387-400.

DONALD A. NORMAN
The psychology of everyday things.
 New York, Basic Books, 1988.

CHRISTINE L. BORGMAN
Why are online catalogs still hard to use ?
 In : *Journal of the American Society for Information Sciences*,
 47 (1996), 7 : 493-503.

DONALD A. NORMAN,
The invisible computer : why good products can fail, the personal computer is so complex, and information appliances are the solution [Reprinted]
 Cambridge, Mass. ; London : The MIT Press, 1999 ; XII,
 302 p. ; ill. ; 23 cm.

CHRISTINE L. BORGMAN
From Gutenberg to the global information infrastructure. Access to information in the networked World.
 Cambridge, Mass., London, The MIT Press, 2000 ; 324 p.
 (Digital Libraries and Electronic Publishing).

Successes and failures of digital libraries. Ed. By Susan Harum and Michael Twindale
 Urbana-Champaign : University of Illinois, Graduate school of library and information science, 2000 ; 134 p.
 (Clinic on library applications of data processing ; 1998).

YOLANDE ESTERMANN ; ALAIN JACQUESSON
Quelle formation pour les bibliothèques numériques ?
 In : *Bulletin des bibliothèques de France*. (À paraître).