

Outils personnels et outils publics, la fin d'une frontière ?

Fondation nationale des sciences politiques

Séminaire *Temps, Médias, Société*

Paris, le 16 décembre 2005

Michel Fingerhut

Lorsqu'on demanda au pianiste hongrois Ervin Nyiregyhazi (1903-1987) pourquoi il n'avait jamais possédé de piano, il répondit : « Est-ce qu'un chauffeur de poids lourds a un camion dans son salon ? » J'ai commencé à faire de l'informatique en 1966 (à cette époque, les ordinateurs avaient effectivement la taille de gros camions), et si ce n'est qu'en 2002 que j'ai acheté un ordinateur personnel (tandis que j'ai dû acheter mon premier livre, à 100 fr., en 1957), cela ne m'a pas empêché de réaliser depuis 1995 quelques « bibliothèques numériques », professionnelles ou personnelles.

J'évacuerai l'aspect purement lexicographique du terme « bibliothèque numérique » : il est évident qu'il ne s'agit pas uniquement de livres, comme il est évident que la plupart des bibliothèques « physiques » – depuis la Bibliothèque nationale jusqu'à celle de l'Ircam – ne comprennent pas que des monographies dans ses fonds. La bibliothèque est, en ce qui me concerne, un organisme social – composé de lieu(x) principalement physique(s), d'objets (matériels et immatériels), de personnes – et fonctionnant selon des modalités et des règlements (tacites ou explicites) visant à cadrer les interactions entre les personnes et l'utilisation des objets pour une finalité particulière.

Pour fixer l'échelle temporelle, je rappellerai que le Web est apparu au début des années 1990, l'internet a été inventé au début des années 1970, tandis que le zéro, qui est à la base du numérique, est attribué aux Babyloniens (III^e siècle avant JC) et l'on a trouvé des tablettes cunéiformes comprenant ce qu'on pense être l'un des premiers catalogues dans le temple d'Enlil à Nippur, datant d'environ 2000 ans avant JC. À cette échelle, nous sommes encore dans les balbutiements de ces technologies dites nouvelles, et il serait futile de tenter d'en prévoir les évolutions à long terme.

En tout cas, l'évolution des techniques de fixation, de production, de stockage et de diffusion de la production de l'esprit humain amène les bibliothèques à remettre en question périodiquement leurs périmètres d'action – et ceci bien avant l'arrivée de l'informatique, des technologies du numérique et de leurs capacités de production, de traitement et de diffusion –, sans pour autant renier leurs missions de base classiques : créer, organiser et préserver des collections, et y fournir l'accès au public.

Ce sont tous ces périmètres qui sont transformés actuellement avec l'arrivée de nouveaux outils de production et d'accès aux « contenus », et donc de nouveaux contenus. Leur apparition est ancrée dans une histoire (celle de la science et des techniques), et vient autant répondre à un besoin d'utilisateurs potentiels qu'à celui de l'inventeur (qui n'est pas le même) ; ils sont donc loin d'être neutres. Avec leur appropriation, apparaissent de nouveaux usages, inattendus pour certains, et donc de nouveaux besoins, parfois contradictoires avec ceux qui leur préexistent.

C'est avec ces préliminaires présents à l'esprit que je vais aborder la question des périmètres :

- Les lieux physiques : depuis l'apparition des catalogues informatisés sur le web, le public peut les consulter à distance. Avec la numérisation de certains documents (sonores, textuels, visuels), il peut aussi accéder à distance à des contenus. Les collections elles-mêmes ne sont plus uniquement localisées dans la bibliothèque, avec la réalisation de catalogues communs à des organismes distincts ou le rajout de « ressources externes » (par exemple : bases de données de périodiques numérisés, sites Web...) au catalogue. Quant aux services de médiation, si certains se faisaient déjà par téléphone, le courrier électronique se banalise et certaines bibliothèques offrent dorénavant un service à distance 24h/24.

- La nature des objets, des collections et leur disponibilité : la numérisation a élargi (tout en la rendant plus floue) la notion de document édité (tout document en ligne est « édité », en quelque sorte), et, de ce fait, gommé des frontières autrefois immuables entre bibliothèques et archives. De toute façon, le public ignore ces frontières, et va chercher les contenus concernant le domaine qui l'intéresse là où ils se trouvent, que ce soit dans une bibliothèque ou dans des archives ; les Canadiens ne s'y sont pas trompés, lorsqu'ils ont fusionné leur bibliothèque nationale et leurs archives nationales¹ en 2004.

La nature de l'objet « livre » lui-même (et celle de l'œuvre en général²) est en train de muter, accompagnée par l'émergence de nouveaux objets. Le livre tel qu'on le connaît se prête facilement à la numérisation³, mais plus difficilement à la lecture continue à l'écran (à l'inverse du disque) : c'est là que l'utilisation d'une prothèse peut se révéler un handicap. À l'inverse, l'hypertexte a fait apparaître de nouveaux « objets » d'écriture qui ne se prêtent pas à la fixation physique sans les dénaturer.

Enfin, entre la disponibilité physique et numérique d'un même ouvrage, le lecteur choisira cette dernière, et s'il ne le trouve pas en ligne, il tentera souvent d'en trouver l'équivalent ou un substitut, ce qui risque de faire tomber en désuétude des pans entiers de collections, ceux qui n'auront pas été numérisés, pour quelque raison que ce soit, et d'encourager la désaffectation des bibliothèques.

- L'organisation de la connaissance : traditionnellement, c'est le bibliothécaire (ou le documentaliste) qui organise les contenus présents dans son rayon d'action, par le moyen de l'indexation qu'il fournit dans les notices du catalogue, que ce soit à l'aide d'un thésaurus standard ou particulier au fonds.

¹ Dont j'avais parlé ici : <<http://mmdl.free.fr/blog-m/?p=203>>.

² Les capacités de reproduction et de découpage à l'infini du numérique, depuis les collages de Max Ernst et la réflexion d'un Walter Benjamin, remettent en question sa notion même.

³ C'est loin d'être le cas pour bien des ouvrages imprimés de grande taille (y inclus les partitions musicales), dont la lecture s'accommode facilement sur support physique et dont l'écran ne fournit qu'une fenêtre forcément réduite.

Or trois évolutions ont vu le jour depuis :

- D'abord, les capacités informatiques à analyser automatiquement les contenus numériques (et pas uniquement textuels) pour fournir des typologies et des cartographies pertinentes, permettant d'y accéder d'une façon non linéaire (l'unidimensionnalité de la présentation des réponses fournies par la plupart des moteurs est sidérante).
- Ensuite, les capacités de certains moteurs de recherche à conserver les requêtes des utilisateurs et à les analyser pour tenter d'en déterminer leur « goût » sans pour autant permettre (encore) de l'explicitier⁴.
- En dernier lieu, le phénomène des *blogs*, qui permettent à tout un chacun de devenir auteur (et créateur, en général) et de s'auto-publier, mais aussi de se créer sa propre taxonomie (phénomène appelé si efficacement en anglais *folksonomies*) pour classer sa production qui reflète finalement sa construction individuelle et collective du sens⁵.

Ce lecteur-auteur-éditeur ne se satisfait plus d'une représentation hiérarchique par disciplines distinctes qui fragmente notre vision du monde ; les réseaux sont passés par là, risquant à l'inverse de mêler irrémédiablement tous les genres.

- Les réseaux sociaux : si le bouche-à-oreille a toujours été un moyen de diffusion de la « connaissance de la connaissance » (par les pairs, par les médias...), la technique a permis de mettre en œuvre des outils pratiques de diffusion de « l'information sociale » à propos de contenus : lorsque l'on consulte la référence d'un ouvrage chez Amazon, on voit quels autres ouvrages ont « intéressé » (la mesure de l'intérêt étant réelle : l'achat) ceux qui ont acheté l'ouvrage en question (on peut y voir aussi quels ouvrages sont cités dans le corps de l'ouvrage en question, autre réseau à ne pas ignorer). Des techniques comme le RSS (syndication informatique) permettent, à l'instar de la DSI⁶ du « passé », de se maintenir au courant, de façon informelle, de ce qui se publie dans des sources choisies intéressantes et pertinentes.

- L'appropriation : traditionnellement, le lecteur pouvait prendre des notes manuscrites (ou sur son portable), recopier des extraits des contenus des ouvrages auxquels il avait accès sur place ; puis il a pu en photocopier certains. L'informatique y a rajouté la capacité de numériser, puis d'accéder rapidement et de naviguer non seulement dans des collections, mais aussi dans des documents individuels complexes (par exemple : de longs enregistrements sonores), selon des structurations diverses et

⁴ Ce qui soulève la grave question de la frontière privé – public, déjà sérieusement remise en question par les méthodes de traçabilité numérique, l'analyse des messages de courrier électronique privé par de tierces parties à des fins commerciales, etc.

⁵ La seule fonctionnalité qu'offrent certains catalogues en ligne est le panier de notices, sorte de taxonomie à deux éléments (rarement plus, et pas hiérarchiques) : *dedans* ou *dehors*.

⁶ Diffusion sélective de l'information, service auquel certains logiciels bibliothéconomiques et documentaires permettent de s'abonner pour recevoir de l'information selon des profils que l'on aura précisé, au fur et à mesure que le système en intègre de nouvelles.

des aides (documentaires ou dérivées automatiquement des contenus - index, résumés...). Elle lui permet de copier des documents ou des parties de document d'un clic, et, ce qui est bien plus important, finalement, de les annoter⁷, de les transformer, de les citer (voire de les plagier) et de les utiliser pour ses propres productions, à l'instar d'un DJ.

• La nature du monde sensible : le savoir ne se constitue pas uniquement par l'entremise d'inventions et ne se nourrit pas uniquement d'informations : il se construit dans le temps, dans l'élaboration d'un système complexe d'accumulation successives, de transformation, d'invention, de production, de validation, d'appropriation et de conservation. L'émergence des moyens d'auto-publication (dont les *blogs* sont l'avatar grand public), de démultiplication à l'infini et de pérennisation *ad hoc* bouleverse et dénature ce processus. Ils favorisent surtout la diffusion rapide de documents de nature plus éphémère (la presse, par exemple), mais aussi celles de plagiats, de falsifications et de rumeurs.

À l'inverse, la conservation organisée à long terme du numérique est une tâche complexe, souvent antinomique aux principes de fonctionnement de l'innovation technique qui nécessite l'obsolescence⁸. La course en avant, de plus en plus rapide, induite par la technique qui s'emballe et les marchés qui en profitent, ne peut souvent s'accommoder d'un patrimoine (culturel, scientifique) plombé par l'immutabilité de sa physicalité. L'innovation est changement, et ce ne sont souvent que ses traces que l'on peut préserver, et c'est l'une des problématiques des évolutions actuelles : que préservera-t-on ? Quels critères en guideront les choix ? Quelle validité auront-ils pour les générations futures ?

Ce sont les subtiles frontières entre expert et amateur⁹, entre information et savoir, entre rationalité et subjectivité, entre individu et société¹⁰, qui sont ainsi remises en

Autre est celui qui est capable de mettre au jour les procédés d'un art, autre est celui qui l'est d'apprécier quel en est le lot de dommage ou d'utilité pour les hommes appelés à s'en servir ! (...) Car cette invention, en dispensant les hommes d'exercer leur mémoire, produira l'oubli dans l'âme de ceux qui en ont acquis la connaissance ; en tant que, confiants dans l'écriture, ils chercheront au dehors, grâce à des caractères étrangers, non point au-dedans et grâce à eux-mêmes, le moyen de s'en ressouvenir ; en conséquence, ce n'est pas pour la mémoire, mais pour la procédure du ressouvenir que tu as trouvé un remède. Quant à la science, c'en est l'illusion, non la réalité, que tu procures à tes élèves : lorsqu'en effet, avec toi, ils auront réussi, sans enseignement, à se pourvoir d'une information abondante, ils se croiront compétents en une quantité de choses, alors qu'ils sont, dans la plupart, incompetents ; insupportables en outre dans leur commerce, parce que, au lieu d'être savants, c'est savants d'illusions qu'ils sont devenus.

Platon,
Phèdre ou de la Beauté

⁷ Et d'annoter aussi les métadonnées (les notices descriptives des ouvrages dans les catalogues en ligne).

⁸ La « destruction créatrice » de Joseph Schumpeter, moteur de l'économie capitaliste.

⁹ Qualificatif que je préfère revendiquer, en ce qui me concerne.

¹⁰ Jean Baubérot parle de « *l'individualisme de masse* ». Cf. Michel Fingerhut : *Livre et liberté*, juillet 2003 (<http://mmdl.free.fr/blog-m/?page_id=198>).

cause, ouvrant les portes à la balkanisation de la connaissance autant (ou encore plus ?) qu'à sa constitution.

• La question des finalités : dans l'annexe à un article universitaire datant de 1998¹¹, les co-fondateurs de Google, Sergey Brin et Larry Page, écrivaient : « *le modèle actuel des moteurs de recherche à vocation commerciale est basé sur la publicité. Le business model de celle-ci ne correspond pas toujours aux critères de choix de contenus de qualité pour l'utilisateur (...). Pour ces raisons et comme le montre l'histoire des médias, il nous semble que les moteurs de recherche financés par de la publicité seront biaisés, de façon inhérente. Il est donc crucial d'assurer l'existence d'un moteur de recherche transparent et situé dans le secteur universitaire.* »

Mutatis mutandis, c'est le meilleur plaidoyer pour une grande bibliothèque numérique européenne qu'on puisse faire, et venant des fondateurs de Google, il n'en a que plus de poids.

Je la verrai, cette bibliothèque, plus proche de ce qu'Amazon met en place que Google (et peut-être pour ce que fera Microsoft avec les fonds de la British Library qu'ils numériseront en 2006) : un dispositif multiculturel, multilingue réparti (des fonds en réseau¹²), intégré¹³, polymorphe, extensible, recomposable et personnalisable, prenant acte de ces évolutions, pour le référencement, la gestion, l'organisation, la circulation et la diffusion de documents de nature différente (texte, image, son..., pour certains numérisés pour d'autres non) et de ressources numériques choisies¹⁴, témoins inaltérés du passé qui se constitue ; contenant des métadonnées de bonne

L'omniprésence de Google impose sa vision. La somme des connaissances est telle qu'elle nécessite des partis pris, explicités ou non : c'est vrai dans le virtuel comme dans le réel, pour les moteurs de recherche comme pour les journaux ou les bibliothèques. Mais les partis pris des moteurs de recherche, dans la sélection et dans la présentation de leurs sources, incluent, à grande échelle, des considérations commerciales (notamment pour ceux qui sont cotés en bourse) et technologiques (sélection des sources, critères de recherche, algorithmes, mesures de pertinence...), qui priment sur le devoir d'information du public ou celui de préservation, de diffusion et de valorisation du patrimoine humain (culturel, scientifique). Un des critères les plus pernicioeux de sélection des sources en est leur popularité ; ce hit parade n'est pas un critère de qualité mais il devient le principal critère de pertinence dans le monde massifié de la mondialisation numérique, où le maître-mot de son darwinisme est la statistique et le chiffre d'affaire.

Michel Fingerhut
Quand Google booste la réflexion
 23 février 2005

¹¹ *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, disponible en ligne à l'adresse <<http://www-db.stanford.edu/~backrub/google.html>> et cité le 8 décembre 2005 par Nicholas G. Carr dans son (excellent) *blog* (<http://www.rougtype.com/archives/2005/12/the_birth_of_go.php>).

¹² Pour éviter la mainmise d'un organisme quelconque sur ces fonds. Google, une entreprise privée américaine, se transforme en fournisseur de contenus numériques, qui ne seront accessibles que par ses propres index. Le danger de leur séquestration, pour des raisons économiques ou politiques, n'est pas à ignorer. On rappellera l'embargo américain décrété en 2004 sur l'édition, à l'encontre de certains pays (cf. <<http://mmdl.free.fr/blog-m/?p=205>>).

¹³ Et non pas une juxtaposition de catalogues, de bases de données, d'annuaires... distincts.

¹⁴ Incluant, par exemple, l'équivalent automatique du récolement des ressources électroniques externes.

qualité¹⁵ ; proposant des moyens de recherche multiples (par index, par texte intégral, par langage naturel, par réseaux sémantiques et sociaux...), intuitifs ou avancés ; permettant à chaque utilisateur de s'en faire « son » catalogue, qu'il pourra renseigner sur la pertinence des réponses fournies, et ainsi l'orienter vers ses propres critères plutôt que ceux du dispositif sous-jacent ; lui offrant les moyens de s'approprier les contenus, de les organiser et de les enrichir ; de communiquer à propos de ces contenus avec d'autres usagers, sur place ou à distance.

Quant à la bibliothèque, elle maintiendra ses missions, mais c'est son fonctionnement qui continuera à évoluer, pour une meilleure collaboration dans la production, la validation, l'organisation et l'utilisation de la connaissance entre le bibliothécaire et l'utilisateur, avec l'assistance de la technique, pour assurer le difficile mais passionnant rôle de médiation, de pôle de référence et d'équilibre entre le local et le global, le privé et le public, le physique et le virtuel, l'instantané et le durable, dans un réseau transdisciplinaire et transculturel en perpétuelle mutation¹⁶.

Dans un entretien accordé récemment au *Magazine littéraire*, le poète et essayiste Nimrod disait :

La perspective, inventée au XVI^e s. dans le domaine de la peinture, viendra hiérarchiser les plans. Mais la perspective n'est pas une vision naturelle, elle est reconstruction. Il faut revenir à Platon pour retrouver toutes ces implications (...), le Platon du Sophiste qui serait du côté de l'art byzantin où tous les plans se chevauchent, bouleversent la perspective et nous avec.

La modernité nous fait paradoxalement changer de perspective et regarder bien loin en arrière pour mieux avancer, et il faut en prendre acte.

Version revue et corrigé d'un texte présenté à la journée d'étude sur l'avenir des systèmes d'informations des bibliothèques, Pau, 25 novembre 2005.

¹⁵ Je ne crois pas à la disparition des métadonnées professionnelles, bien au contraire (de même que ce n'est pas parce que la conduite d'une voiture se simplifie que le moteur disparaît). Mais il est d'autant plus nécessaire d'adapter leur format pour permettre une description plus aisée des « nouveaux documents » et de leur gestion, d'y inclure de nouvelles fonctionnalités (la gestion numérique des droits, par exemple) et d'automatiser, autant que faire se peut, leur mise à disposition (pour éviter d'avoir à cataloguer).

¹⁶ Si on dit du monde qu'il se bibliothécarise, il ne faut pas oublier que le bibliothécaire se spécialise lui aussi.