

La Bibliothèque virtuelle ou de l'ars memoria à Xanadu

Yannick Maignien

► **To cite this version:**

Yannick Maignien. La Bibliothèque virtuelle ou de l'ars memoria à Xanadu. Bulletin des bibliothèques de France, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB), 1995. <sic_00000301>

HAL Id: sic_00000301

https://archivesic.ccsd.cnrs.fr/sic_00000301

Submitted on 20 Dec 2002

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Yannick Maignien

Bibliothèque nationale de France

Mission Politique scientifique de numérisation

LA BIBLIOTHEQUE VIRTUELLE

ou de l'Ars memoria à Xanadu

La numérisation des documents entraîne avec elle une nouvelle conception de l'encyclopédisme. Le rapport du lecteur à de grands volumes d'informations (par exemple, des collections patrimoniales) est en profonde mutation. Ce rapport renvoie à une organisation réticulaire et complexe du savoir enregistré. Dès lors que de nouveaux acteurs entrent en jeu sur les réseaux, de nouveaux outils de travail, de sélection et de navigation doivent être pris en compte pour exploiter cet encyclopédisme réticulaire.

Opérer une meilleure visibilité sur de grands corpus de connaissances, tel est ce que permet l'ordinateur : une grande recension de textes, *via* l'indexation, l'annotation, le classement de ces connaissances et les possibilités de réactiver de façon automatique le capital d'un travail entrepris de longue haleine.

Constituer une bibliothèque immatérielle n'est donc ni un acte neutre, ni un processus dans la continuité de la bibliothèque physique¹.

Une tradition ancienne, une conjoncture nouvelle

Cet article voudrait montrer que cette entreprise liée aux nouvelles technologies d'information et de communication renoue avec une autre entreprise plus fondatrice, plus fondamentale de constitution encyclopédique du savoir, ce qui suppose que l'on ait un regard réflexif sur le rôle de l'imprimé et du livre. Une bibliothèque suppose toujours une conception organisatrice, donc une vision encyclopédique, un projet de rassemblement et de mise en "circulation" des savoirs. C'est ce double projet qui est profondément modifié par le média électronique : possibilités nouvelles de lecture, de recherche, mais aussi contraintes nouvelles pour l'ensemble de la chaîne éditoriale.

La bibliothèque virtuelle devra à l'avenir redéfinir ses missions au sein de cette mutation de la chaîne éditoriale : auteurs, éditeurs, producteurs d'outils informatiques de structuration (à la place des imprimeurs), bibliothèques et circuits de librairie, lecteurs ou nouveaux usagers des ressources comme l'école ou l'université, devenant eux-mêmes utilisateurs-producteurs d'informations.

Une nouvelle économie de l'information s'ouvre : encore faut-il pleinement en mesurer les effets si nous ne voulons pas manquer les occasions qui se présentent.

¹ Un excellent article de Pierre-Marie Belbenoit-Avich a déjà présenté l'essentiel de cette mutation en matière d'organisation bibliothéconomique. Cf. "La bibliothèque électronique", Bulletin des bibliothèques de France, t. 38, n° 6, 1993, p. 60-64.

Les réflexions suivantes, en marge des missions de la politique de numérisation de la Bibliothèque nationale de France, n'engagent – comme l'on dit –, que leur auteur, tant les visées prospectives qui l'inspirent relèvent de domaines largement ouverts.

La numérisation des documents (textes, images fixes et animées, sons, manuscrits, collections du patrimoine, données scientifiques), par la dématérialisation des supports qu'elle implique, ne saurait offrir uniquement des objets de substitution, simplement plus transmissibles ou réitérables à volonté. En réalité, c'est la question du virtuel qui est en jeu, c'est-à-dire de l'espace heuristique de simulation, au sens d'illusion et de création mi-sensible mi-intelligible, par lequel un nouvel espace de connaissance serait possible.

Celui-ci permettrait de mieux penser la complexité, la diversité du monde et des phénomènes tels que les sciences humaines ou de la nature les construisent et tentent de les appréhender. Les strictes représentations “positivistes” ou objectives pouvaient se satisfaire de traités monographiques spécialisés ; la science mouvante et complexe du monde contemporain nécessiterait des approches multiples par construction d'univers imaginaires.

Encore faut-il que la bibliothèque ne fasse pas fausse route dans sa perception du rôle de la numérisation, c'est-à-dire de ses pouvoirs et de ses limites.

De Raymond Lulle à Ted Nelson

En fonction des technologies nouvelles de stockage, de traitement et de diffusion de l'information que permet la numérisation des documents et l'informatique, il y a lieu de se poser les questions que, par exemple, la culture de la Renaissance se posait autour des techniques des arts de la mémoire, pour conserver, mais aussi organiser et rendre fécond le rassemblement du savoir. Cette question, vive à nouveau grâce notamment à l'informatique, d'un *ars inveniendi*, dessine les contours d'un nouvel encyclopédisme.

Regardons cette image énonçant en une page l'essentiel de la philosophie de Raymond Lulle. Elle permet à la fois de mémoriser, d'organiser, de combiner et d'inventer. En 1945, cinq siècles plus tard, Vannevar Bush², tirant toutes les conclusions de la prodigieuse miniaturisation et de la mécanisation photoélectrique, développait sous le nom de Memex (*Memory Extender*) cette même notion d'une bibliothèque universelle de documents, fondée sur l'idée d'association et de liens (contre l'indexation), de construction de bibliothèques personnelles de *trails* (de “pistes”). À l'aube de l'informatique, était déjà posée la question de l'optimisation du rapport entre masse d'informations, miniaturisation des supports, rapidité de la recherche, et intelligence des liens associatifs, dans la lignée des Leibniz et Babbage.

Ted Nelson, avec son projet *Xanadu* dans les années 80, reprend la réflexion de Vannevar Bush, indépendamment des techniques – analogiques là, ici devenues électroniques pour systématiser une combinatoire hypertexte de tous documents numériques (image, texte, sons...) au sein du *Literary Machines*. Toutes proportions gardées, sont reposées les questions de la possibilité d'accéder à des données complexes avec des nouveaux impératifs de diffusion, d'ouverture des connaissances, comme le firent les hommes de la Renaissance avec l'imprimé.

² Vannevar BUSH, “As we may think”, The Atlantic Monthly, 1945 ; Theodor H. NELSON, *Literary Machines*, Sausalito, Mindful press, 1990.

L'encyclopédisme, c'est d'abord cette liberté de pouvoir disposer du passé, l'oublier ou le réactiver autant que de besoin, au risque sinon d'osciller entre le vide de l'ignorance, ou le trop-plein de l'encombrement hétéroclite.

Cela a souvent été relevé, la révolution numérique actuelle n'a de référent comparable que le passage du manuscrit à l'imprimé au XV^e siècle, à l'imprimé comme nouveau moyen de stockage, d'organisation et de circulation des savoirs.

Sans trop tordre la réalité historique, l'image proposée ici, symbole de l'*Ars combinatoria*, une “ illustration ” de l'édition de Valence du *Liber de ascensu et descensu intellectus* de Raymond Lulle (1512) est de façon troublante un présage de l'hypertexte : les liens entre notions et icônes ont un rôle de concentration et d'organisation des connaissances, au profit d'une technique de la mémoire, par disposition de lieux et parcours imaginaires. Ces techniques ont ensuite été chassées par le savoir typographique et la disjonction du texte et de l'illustration, au point d'être oubliées jusqu'à leur récente redécouverte³.

Tout y est, l'image synthétique et symbolique de l'organisation du savoir, l'escalier de l'ascension et de la descente des différents niveaux de l'être, des choses à Dieu, en passant par le minéral, le végétal, l'animal, l'humain, les anges. Les liens ou “ boutons de liaison ” attachent des icônes à ces notions. Mais c'est aussi une représentation de l'organisation dynamique du savoir, avec la roue de la Combinatoire des neuf “ dignités ” selon les notations lulliennes : B, C, D, E, F, G, H, I, K.

Comme Giordano Bruno, autre grand Renaissant des arts de la mémoire, l'homme contemporain doit systématiser le savoir ancien pour l'obliger à produire du nouveau⁴.

De la Kabbale à Leibniz, et jusqu'à Borges, l'espoir est mis dans l'économie combinatoire que permet la finitude alphabétique (dans la *Restitution universelle*, Leibniz place dans la combinatoire infinie des 26 lettres l'origine de la bibliothèque universelle de tous les livres qui existent et ceux qui n'existent pas. L'idée d'exhaustivité ouvre sur l'infinité du savoir. La sélection des rares bons ouvrages ne peut se faire que sur fond de l'existence des innombrables mauvais. Partiellement, l'ordinateur reprend cette visée dialectique de la finitude et de l'infini des grands nombres, de l'efficacité de la sélection singulière au moment où la bibliothèque avoue son impossible exhaustivité.

L'imprimé n'a-t-il été qu'une parenthèse, au sens où, pendant des siècles, se serait développé le plus fascinant système de conservation et de circulation des connaissances ? L'invention des caractères mobiles d'imprimerie aurait surtout été la révolution éditoriale de masse, raffinée, à peu de frais et où la rationalisation des processus de fabrication a peu à peu été conquise par l'informatique.

³ Bernard Cerquiglini note, à propos du texte médiéval et du travail nouveau que permet l'ordinateur : “ L'informatique, qui mine sourdement la pensée textuaire, peut nous être d'une aide technique digne de réflexion. L'écrit écranique est à la fois dialogique (interaction du consultant et de la machine) et spatial (consultation conjointe, par fenêtrage, de données appartenant à des ensembles disjoints). (...) La page serait alors tournée. L'écrit électronique, par sa mobilité, reproduit l'œuvre dans sa variance même. L'informatique retrouvant, en deçà de la Modernité, le chemin d'une ancienne littérature dont l'imprimerie avait effacé la trace : c'est un beau sujet de méditation ”.

“ Variantes d'auteur et variance de copiste ”, La naissance du texte, Paris, J. Corti, 1989.

⁴ Cf. Paolo ROSSI, *Clavis Universalis*, Grenoble, J. Millon, 1993 ; Frances YATES, *L'Art de la mémoire*, Paris, Gallimard, 1975, 440 p. ; Jacques ROUBAUD, *L'invention du fils de Leoprepes*, Paris, Circé, 1993, 168 p.

En un sens, l'histoire du livre se referme, et avec elle une tradition bibliothéconomique très spécifique. S'ouvrent d'autres vocations, d'autres missions mnémotechniques inédites, sans que le programme précis puisse en être établi.

Le numérique n'est pas que le support de substitution, la reproduction d'un objet matériel, textuel, sonore ou iconique, c'est une "autre" machine éditoriale, non encore développée, c'est un "autre" agencement des relations aux connaissances et aux autres, une autre communauté en gestation.

Pour autant, ne concluons pas que l'informatique est, déjà, cette nouvelle machine mnémonique pouvant exhumer les savoirs à la demande et mettre en relation facilement les compétences dispersées. Cela ne se fera que selon des règles qui sont encore largement à établir.

Le présent article essaie de montrer qu'un certain nombre de difficultés ne sont nullement d'ordre technique ou informatique, mais culturelles ou théoriques, sinon politiques.

L'impossible encyclopédisme

Avant de préciser ces points, accordons que l'acte de connaître ne se satisfait plus d'une visée classificatoire : la fécondité des connaissances est plus attendue dans les limites, les marges, transdisciplinarités, interdisciplinarités, multidisciplinarités, hybridations, métissages, et autres rapprochements inattendus (Edgar Morin met particulièrement en valeur ce type d'encyclopédisme).

Complexité des savoirs, nouvelles alliances entre sciences "dures" et sciences humaines sont partout de mise : indéterminisme, impossibilité de dissocier les conditions d'une connaissance des résultats de la connaissance elle-même, ambivalence des approches, comme si l'instabilité des savoirs était essentielle, inhérente désormais au rapport de l'homme à son milieu. Au-delà des mutations du média, c'est aussi la scientificité qui est en question, nécessitant de nouvelles formes d'exposition, de cohérence, d'interactivité, de régulation ou d'effectivité au sein des communautés savantes.

Une maîtrise synthétique des savoirs, sinon par des approches collectives, n'est peut-être plus de mise actuellement (l'était-elle déjà au début du siècle ?) devant l'explosion des sciences hyperspécialisées.

Les grands scientifiques du début du siècle avaient encore l'idéal scientifique que la vérité doit être énoncée en loi. " *Le chaos n'a pas d'histoire, pas plus qu'il n'a de loi* ", dit Henri Berr dans la *Synthèse en histoire*⁵ Nous savons au contraire que nous devons compter avec lui, intégrer l'instabilité, l'incertitude, la complexité et que rien n'inscrit nécessairement le changement dans un "développement". Aussi, les données encyclopédiques actuelles requièrent-elles des précautions, des constructions de méthodologie précaires et révocables.

Au-delà de cette épistémologie de la complexité – ce dont traitent les sciences –, c'est le discours, le langage qui instaure ces connaissances en tant que sciences, qui est suspect. Le passé culturel lui-même ne cesse d'être revisité, réinterprété, réagencé dans de nouvelles *épistémè*. Tout savoir est l'objet d'une volonté de vérité qui est suspect, du moins dont le mobile demande à être justifié, explicité, décrit.

C'est l'apport de Michel Foucault d'avoir conduit cette réinterprétation des savoirs, dévoilé les *épistémè*, les généalogies, l'archéologie qui constituent ces savoirs, révélé la diversité des approches de "jugement". En ce sens, il ne peut être fait abstraction d'une herméneutique ou d'une

⁵ Henri BERR, *La Synthèse en histoire*, Paris, F. Alcan, 1911.

phénoménologie, en tout cas des divers niveaux de complexité où est interrogé un ensemble de données dans des systèmes préétablis⁶.

C'est patent dans sa leçon inaugurale au Collège de France, *L'Ordre du discours*, dès 1970, et sans doute dès *Les Mots et les choses* où se pose la question – comme chez Borges – de la classification du savoir, de la taxonomie, et dès *L'Histoire de la folie à l'âge classique*, où le rôle prescriptif de pouvoir et d'interdit du langage est déjà en jeu.

Si, en effet, des recherches aidées par l'informatique sur des grandes masses de textes sont rendues à la fois nécessaires et possibles, c'est bien parce que, dans le sillage de Foucault, et sans doute sans le reconnaître, on recherche des “volontés de savoir” propres à telle époque, à des rapprochements de disciplines ou de documents qui n'ont de sens que dans telle *épistémè* précise. “*Quelle est, à travers nos discours, cette volonté de vérité qui a traversé tant de siècles de notre histoire ?*”, interroge-t-il dans cette leçon inaugurale, posant ainsi une question renouvelée de constitution du savoir, une question de fondement encyclopédique.

Les notions d' “auteur”, d' “oeuvre”, l'organisation et l'émergence des “disciplines” sont interrogées, dénoncées, resituées.

“*La discipline est un principe de contrôle de la production du discours*”, (*L'Ordre du discours*, p. 37). “*A l'intérieur de ses limites, chaque discipline reconnaît des propositions vraies et fausses ; mais elle repousse, de l'autre côté de ses marges, toute une tératologie du savoir*”. Et celle-ci est toujours à réinterpréter, à réinterroger, dans la mesure où ces “monstres rôdent dont la forme change avec l'histoire du savoir” (p. 35), “et les monstres vrais peuvent être temporairement en marge (comme Mendel pour la découverte de la génétique), en avance sur des auteurs qui sont “dans le vrai” de leur époque et qui ne sont qu'erreur disciplinée”. Il s'agit de “traiter, non pas les représentations qu'il y a derrière les discours, mais des discours comme des séries régulières et distinctes d'événements”.

Le nouvel encyclopédisme requis est donc d'essence : c'est le discours comme agencement des disciplines, dépôt des textes écrits, édités, qui n'est plus une image fiable, confiante, d'un réel supposé. Aussi n'y a-t-il d'issue que dans une large et infinie confrontation des sources, des péri-textes, des contextes, des références.

C'est bien ce débordement monstrueux, erratique de savoirs qui nécessite de nouvelles méthodes et un nouveau média.

Ordonnateur, ordinateur de textes, l'informatique permet, autorise presque le démembrement des volumes, reliures, séries, collections, pour s'ouvrir, à partir de ces matériaux, sur une recomposition, de nouveaux liens, de nouvelles circulations hypertextuelles, créant ainsi des éditions ou des bibliothèques virtuelles, faisant apparaître des généalogies, des archéologies, des volontés de savoir non directement explicites, libres des organisations de savoir qui furent les leurs.

Ce n'est sans doute pas simple coïncidence de faire l'hypothèse de ce rapprochement, dès lors que l'ordinateur et la numérisation sont des “machines” à structurer le savoir, comme l'ont été le *codex* ou les rouleaux d'Alexandrie.

L'espace-temps informationnel

⁶ Cf. Michel FOUCAULT, *Les Mots et les choses : une archéologie des sciences humaines*, Paris, Gallimard, 1966, 400 p., et *L'Ordre du discours : leçon inaugurale au Collège de France*, prononcée le 2 décembre 1970, Paris, Gallimard, 1971, 84 p. (Collection blanche).

Le texte est devenu matière à “ traitement ”, non situable sans une cartographie du savoir devenu mobile, précaire.

*“ La bibliothéconomie moderne, pour s'être installée en ces lieux désertés par les philosophes depuis la fin du XIXe siècle, a rencontré un tel problème : comment classer les connaissances ? Comment construire une bonne systématique du savoir ? Les rangements arbitraires des classifications décimales (Dewey, CDU), avec leurs ramifications indéfinies, parviennent bien à maîtriser les développements exponentiels de la collection des connaissances mais se heurtent au problème des interprétations multiples et des inter-relations : les champs cognitifs se superposent à l'infini, et les objets se fragmentent et se réordonnent différemment avec le temps. Comment maîtriser cette mouvance ? Peut-on, du reste, la maîtriser ? ”*⁷ demande Daniel Parrochia, et de répondre dans le sens de Pierre-Marie Belbenoit-Avich : *“ La fonction principale de la bibliothèque est selon nous, d'aide à la détermination autant qu'à la recherche. Elle concerne l'avenir, non pas le passé. Les moyens modernes de traitement de l'information démontrent avec éclat qu'une telle finalité est devenue nécessité ”*⁸.

Daniel Parrochia rappelle à juste titre que les formes du savoir sont coextensives des techniques d'organisation de ce savoir. Les types de rassemblement du savoir ont été successivement : 1. séquentiel linéaire (listes) ; 2. central circulaire (organisation encyclopédique-systématique) ; 3. relationnel-réticulé, pour l'époque contemporaine.

Les techniques d'information et de communication numériques conduisent donc à un type d'encyclopédisme “ relationnel-réticulé ”, dont nous essayons de préciser ici le type de virtualité qu'à la fois il implique et suppose.

“ La réalité ”, dit Ted Nelson “ est ce qui est, la virtualité est ce qui semble être ”⁹ et il précise le sens de cette “ simulation ” sur laquelle repose le virtuel : “ la virtualité a deux parties : une structure conceptuelle et le sensible ” (Virtuality has two parts : conceptual structure and feel).

La bibliothèque virtuelle signifie donc que cette “ *circularité et circulation du savoir* ” ne peuvent être atteintes qu'au terme d'une nouvelle structure conceptuelle et de représentation sensible.

La numérisation comme dématérialisation

Le passage au virtuel est ce mixte inédit, sensible et intelligible, créant un espace de simulation, intégrant une dimension imaginaire, au sens à la fois d'illusion, mais aussi de création.

Pour bien comprendre la révolution en cours, soulignons, si besoin était, que celle-ci consiste à prendre possession du langage par le codage alphanumérique, à prendre possession de l'image par la “ digitalisation ” en mode graphique, (en description par pixel, assortie des techniques de compression du signal, donc dans un système icono-numérique très singulier), ajoutons en troisième lieu la prise de possession du calcul, de la computation logique, de la combinatoire par le système binaire, (mais c'est l'origine même de l'entreprise de John Von Neumann¹⁰). Mise à profit dans les scénarios prédictifs, cette combinatoire peut se développer en intelligence artificielle, en systèmes

⁷ Daniel PARROCHIA, Philosophie des réseaux, Paris, PUF, 1993, p. 185.

⁸ Id., p. 189.

⁹ Theodor H. NELSON, *ibid.*, p. 27.

¹⁰ John von NEUMANN, L'Ordinateur et le cerveau, prés. et trad. de Dominique Pignon, Paris, La Découverte, 1992, 132 p. J. Von Neumann (1903-1957) est un des pionniers de l'informatique.

experts, en intégrant de larges pans de comportements hypothético-déductifs, de démarches cognitives.

Il faut bien garder à l'esprit ce qu'ont d'hétérogènes ces différentes prises de possession du "réel" par le numérique (*cf.* par exemple les oppositions entre mode image et mode texte en ce qui concerne un texte). L'informatique n'a pas encore gagné sa dignité sémantique qui consisterait à articuler ces trois niveaux spécifiques de communication. Cela touche à la racine même de ce qui constitue l'écriture et la signification : cela tient à des notions aussi fondamentales que celles de l'arbitraire du signe, de la convention, de l'iconicité, de la logique ; et donc à un nouvel agencement des notions que Peirce a par exemple mises en avant dans sa sémiotique : *représentamen*, objet et interprétant¹¹.

Ted Nelson critique la pauvreté de l'informatique bureautique à laquelle nous sommes parvenus : représentation de bureau avec dossiers et documents qui s'amoncellent, se recouvrent, fenêtres qui se chevauchent, icônes multiples¹²... Et Ted Nelson de rappeler combien sont beaucoup plus riches les techniques de montage et les scénarios cinématographiques, et même les jeux électroniques pour enfants.

De ce point de vue, effectivement, la numérisation et l'informatique sont encore dans les limbes de l'archaïsme, d'un technicisme qui n'a pas encore développé son art, tant que des métalangues ne pourront proposer des aides visuelles, sonores ou textuelles autrement adaptées et souples. Il s'agit en fait de cultiver la mise en scène (le *design*) des connexions.

De fait, il n'y a pas encore réellement cohérence des traitements numériques, en tout cas au-delà de l'illusion représentative opérée inconsciemment par l'observateur (lecteur, auditeur, joueur, spectateur...) qui est seul à même d'opérer la synthèse cognitive.

On pourrait dire que la bibliothèque n'est pas directement concernée par les propos qui précèdent. En un sens, la Bibliothèque nationale, par exemple, enregistre par dépôt légal cette avalanche de "produits", progiciels, programmes, multimédias, jeux, enregistrements numériques (CD-Rom, etc.) ; mais, sauf à abdiquer d'emblée et pour longtemps la mission qui a toujours été la sienne de constituer les conditions de l'encyclopédisme, la bibliothèque ne saurait être indifférente à cette impasse dans laquelle la place l'informatique dans son état actuel.

Vers une architecture numérique

L'informatisation des bibliothèques a poussé très loin la sophistication des catalogues et de l'ensemble des techniques périphériques à cette normalisation du signalement bibliographique. D'autres fonctions récurrentes et invariantes peuvent faire aussi l'objet d'automatisation et d'assistance informatisées. Par ailleurs, la numérisation des fonds (textuels, iconiques, sonores, etc.) est entreprise en fonction de problématiques de conservation : substitution à l'original, protection des fonds, prêt, service au lecteur, possibilités statistiques, éventuellement télécommunication des documents, mais toujours en référence à la logique de l'imprimé.

On le sait, une des originalités fortes du projet de la Bibliothèque de France était d'avoir conçu les prototypes de station de travail sur les documents numérisés, parallèlement à l'entreprise de numérisation des collections.

¹¹ Charles PEIRCE, *Ecrits sur le signe* ; rassemblés, trad. et commentés par G. Delédalle, Paris, Le Seuil, 1978, 272 p. (l'Ordre philosophique).

¹² Theodor H. NELSON, "The right way to think about software design", Autodesk, Sausalito, Mindful Press, 1989.

À partir d'un accès au catalogue informatisé, le lecteur choisit son corpus personnel par téléchargement depuis le serveur. Il structure ensuite sa bibliothèque personnelle, à l'aide d'outils de surlignage, d'annotation, de marquage, de repérage de chemins de lecture, d'unités logiques de découpage, outils homogènes et facilement formatables en fonction d'une recherche donnée et permettant au lecteur de traiter une grande masse de documents au sein du même espace électronique.

Les sous-ensembles qu'il souhaite obtenir en mode texte, s'ils ne le sont déjà au sein de la collection, le sont facilement par reconnaissance optique de caractères des fichiers concernés. Des recherches d'ordre lexical ou linguistique sont dès lors possibles et viennent compléter les outils de structuration de son corpus.

Le lecteur peut aussi alimenter ce corpus par les bases de données ou les CD-Rom consultables en réseau. Il peut aussi alimenter cette bibliothèque personnelle par ses propres documents. Réciproquement, il peut conserver d'une session à l'autre l'avancée de son travail, le dupliquer et l'exporter en fin de session.

Le PLAO (poste de lecture assistée par ordinateur) est un *scriptorium* électronique, permettant de mettre en mémoire, réactiver et organiser le travail de lecture et d'écriture.

Le travail de lecture savante est analysé en fonctions génériques pour être intégré dans des automatismes d'assistance, afin d'offrir à l'étude des possibilités sans égales de mémorisation, d'organisation et de vitesse de réactivation des données.

Toute cette conception de la lecture savante a été, on le sait, élaborée par Jean Gattégno, Alain Giffard, Bernard Stiegler, Jacques Virbel¹³ et de nombreux participants associés en tant que “ grands lecteurs ”. La gageure n'était ici pas seulement ni essentiellement technique. Elle supposait d'abord d'avoir une claire vision des possibilités et des limites attendues de l'assistance informatique. Le poste de lecture assistée par ordinateur, par l'acuité de la méthode de conception à partir des usages, a montré ce qu'on pouvait attendre de l'informatique de façon prémonitoire, puisque, quatre ou cinq ans après, ces fonctions tendent à se retrouver de façon banalisée sur le marché des logiciels.

Si la numérisation des fonds (largement engagée au sein de la Bibliothèque nationale de France, par le programme des 100 000 volumes et des 300 000 images numérisés), et la conception de postes de lecture assistée par ordinateur vont de pair, l'informatisation du catalogue, au sens classique du terme, est loin de pouvoir assurer la cohérence de l'architecture à venir du système virtuel que nous évoquions plus haut. Pourquoi ?

De nouveaux modes de recherche

Pour répondre à cette interrogation, posons-nous la question la plus originaire qui soit : que veut dire chercher, “ retrouver ” un ouvrage ? Question encyclopédique s'il en est. Re-trouver ce que l'on connaît déjà ? ou au contraire, ou par ailleurs, trouver, sauver de l'oubli, sans le connaître, un document dont on a besoin dans le cadre d'une recherche ? C'est bien sûr ce second cas de figure qui est intéressant : l'instruction de la stratégie de recherche heuristique est l'une des questions majeures du projet de numérisation de la BNF. “ *Savoir ce que l'on cherche, c'est comprendre ce que l'on trouve* ”, disait Bachelard, mais on sait aussi, depuis, que ce savoir préalable est un horizon de complexité, un spectre de diversité et de sources multidisciplinaires, et que le résultat de la requête

¹³ Cf. Jacques Virbel, “ La lecture assistée par ordinateur et la station de lecture de la Bibliothèque de France ”, Les banques de données littéraires, Limoges, PULIM, 1993.
Jacques VIRBEL, “ Annotation dynamique et lecture expérimentale : vers une nouvelle glose ? ”, Littérature, n° 96, d'été 1994.

peut être aléatoire, non déterministe, et pourtant compréhensible. L'adage peut se renverser : “ *Comprendre ce que l'on trouve, c'est savoir ce que l'on cherche* ”.

Terry Winograd, à ce sujet, note¹⁴ : “ Toute ouverture de nouvelles possibilités en ferme d'autres, et ceci est particulièrement vrai avec l'introduction de nouvelles technologies. Considérons par exemple la possibilité d'une “ bibliothèque électronique ”, dans laquelle on peut rechercher des items selon des techniques sophistiquées de catalogues basées sur la publication (auteur, éditeur, titre) et le contenu (tel que les catégories de la Bibliothèque du Congrès ou le système des mots clés). Si nous acceptons le champ produit par ces classifications comme pertinent pour trouver des livres, le système est approprié et utilisable. Toutefois ce n'est pas forcément le bon choix. Cette méthode permet au lecteur de trouver un livre dans un endroit spécifique, mais il lui interdit de “ feuilleter ” au hasard parmi les étagères. Si on admet l'importance du contexte et du “ pro-jet ”, il est clair que les rencontres inattendues peuvent avoir plus d'importance que les recherches précises. Si le problème est “ trouver un livre et une information précise ”, alors, le système est bon. Mais si le contexte est “ trouver des livres utiles pour ce que nous sommes en train de faire ”, il est moins efficace, parce que l'efficacité ne peut pas être facilement formalisée. En créant un outil, nous changeons la façon dont les gens utilisent la bibliothèque et ses matériaux ”.

L'organisation des disciplines et le classement bibliothéconomique sont dans le droit fil des documents papier, reliés, sériés, identifiés et décrits dans des notices bibliographiques.

Dans une bibliothèque électronique, l'ouvrage et son classement deviennent deux choses indépendantes, l'ouvrage appartient à un fonds indifférencié, du moins où les mobiles et les stratégies d'acquisition ont disparu. Par contre, le lecteur doit pouvoir disposer d'instruments pour instruire sa requête (hors le cas encore une fois où il sait quel livre demander), pour “ formaliser l'efficacité ” de ce type d'agrégation de connaissances autrement féconde dont parle Fernando Flores.

Aussi, à charger en aveugle, dans le cours de cette enquête évolutive, de nombreuses sources non déjà repérées, le lecteur est devant son ordinateur comme devant une boîte noire, un écran sans profondeur. Il lui faut recréer cette profondeur virtuelle, faite d'informations intelligibles, mais aussi de formes sensibles des représentations, les mettant en perspective en fonction des contraintes de l'enquête.

L'idée de navigation

L'instruction de la recherche, sa navigation, comme on dit maintenant (ou comme le disait déjà Leibniz), suppose une cartographie, même sommaire, permettant au lecteur l'exploration sur ces mers inconnues. Et en mer ouverte, c'est d'abord son propre esquif qu'il faut situer ! De fait, c'est toute une hiérarchie d'aides documentaires qui peut être convoquée, mais ces différentes couches d'informations intermédiaires doivent elles-mêmes être organisées, appréhendables dans un espace, une visibilité, une synthèse, même provisoire, précaire (une classification mobile, dirait Leibniz), qui sauve de la dispersion, de la perte de sens.

C'était déjà une des questions centrales des techniques et des arts de la mémoire, avant la Renaissance, où la visibilité synthétique associe image, lieux et notions liés dans des arbres ou des cercles du savoir.

C'était déjà la question cruciale que se posait Leibniz (*La restitution universelle*) dans l'*Horizon de la doctrine humaine*, réflexion sur la finitude de la langue, sur la marque comme signe mémoriel :

¹⁴ Terry WINOGRAD ; Fernando FLORES, *L'intelligence artificielle en question*, Paris, PUF, 1989, 296 p. (La Politique éclatée).

“ *Les mots ne sont pas seulement des signes de ma pensée pour les autres, mais aussi des marques (Notae) de ma pensée pour moi-même* ”¹⁵.

Il s'agit bien en effet non seulement de situer objectivement une source, un document, mais aussi d'indiquer “ la route sur le fonds ”, le sens subjectif (le sujet pouvant être collectif), que prend l'instruction d'une recherche, en fonction de buts qui restent à atteindre, de continents à découvrir. L'aide documentaire relève d'une assistance “ embarquée ”, nécessité interactive qui sera encore plus impérieuse lorsque les ressources numérisées seront elles-mêmes dynamiques, cinétiques (images et fonds audiovisuels animés).

On sait que c'est une des questions majeures des recherches documentaires, dans le domaine de l'hypertexte et de la structuration, du marquage, non pas seulement des formes logiques des contenus (par exemple avec les DTD-Document Type Description SGML¹⁶) mais aussi des balisages des formes logiques des étapes de recherche elles-mêmes (comme par exemple avec les logiciels *HyTime*), et ce, dans tout le spectre des textes, des images fixes et animées, donc de tout objet numérique.

Les notions, qui restent à formaliser et à développer de façon informatique, de cinécartes¹⁷ ainsi que de cinéthèques¹⁸, vont dans le sens de cette cartographie dynamique au centre des entreprises de repérage documentaire.

Aux Etats-Unis, notamment sur ce sujet, un grand programme de coopération sur le thème des *Digital Libraries* est soutenu par la NSF (*National Science Foundation*), la NASA, l'ARPA (*Advanced Research Projects Agency*) du département de la Défense, et regroupe six grandes bibliothèques universitaires.

Des programmes communautaires européens, à une moindre échelle, vont dans ce sens. La Bibliothèque nationale de France a engagé un programme communautaire dénommé Memoria sur ce sujet avec Oxford Texts Archives, le Centre de recherche linguistique de Pise, l'IRIT (Institut de recherche en informatique de Toulouse), les sociétés informatiques AIS-Advanced Information System Berger-Levrault et Cap Gemini Innovation.

*Bernard Stiegler souligne cet aspect dans un article récent : “ La navigation dans un grand système d'information électronique est l'un des grands problèmes de l'information documentaire dans la mesure où il s'agit d'un espace virtuel dont il n'est pas possible d'avoir une vue d'ensemble. Or, l'inscription dans la base par l'utilisateur de ses propres parcours et, plus généralement, le développement des possibilités d'appropriation personnelles du contenu et de l'organisation est sans doute le meilleur moyen pour transformer l'espace documentaire virtuel en espace documentaire habité*¹⁹”.

¹⁵ Gottfried Wilhelm LEIBNIZ, *La restitution universelle. De l'horizon de la doctrine humaine* ; trad. de Michel Fichant, Paris, Vrin, 1991, 218 p. (Bibliothèque des textes philosophiques).

¹⁶ Standard Generalized Markup Language, dont notamment l'entreprise internationale de la Text Encoding Initiative.

¹⁷ Pierre LEVY, *L'intelligence collective : pour une anthropologie du cyberspace*, Paris, La Découverte, 1994, p. 183.

¹⁸ Michel SERRES, *Atlas*, Paris, Julliard, 1994.

¹⁹ Cf. Bernard Stiegler, “ Machines à écrire et matières à penser ”, *Genesis*, n° 5, 1994, Paris, J.-M. Place ; Jean-Louis Lebrave, “ Hypertextes, mémoires, écriture ”, *ibid.*

Le *Literary System* de Ted Nelson insiste sur cette dimension évolutive des versions à garder en mémoire dans la structuration hypertexte des documents, sur la nécessité de disposer d'outils de repérage, de représentation à la fois synthétiques et pouvant être explicités autant que de besoin.

Le véritable sens de l'hypertexte et l'hyperdocument serait donc la façon contemporaine de mettre en "ordre" de telles complexités, dans la mesure où ce qui importe, c'est d'introduire de l'organisation dans les classes de relations qu'entretiennent entre elles des données, plus d'ailleurs que de seulement classer les données elles-mêmes. La structure d'une base de données n'est donc autre que celle de la communauté savante qui produit et utilise ces données, à une période spécifique.

Créer des espaces heuristiques virtuels

Il s'agit bien de créer des espaces heuristiques virtuels, qui se déforment et s'adaptent aux recherches menées, qui s'enrichissent ou se simplifient en fonction des étapes temporelles de la recherche et des exigences d'exhaustivité ou de singularité voulues par le chercheur.

Pour l'heure, cette cartographie documentaire et heuristique dynamique n'est qu'une vue de l'esprit en matière d'assistance par ordinateur et les "routeurs" sont pour longtemps encore les conservateurs ou bibliothécaires des salles de catalogues connaissant leurs fonds... et leurs lecteurs. Mais tout mène à penser que cette question, au centre des grands enjeux des sciences de l'information et de la gestion électronique de documents fera bénéficier les bibliothèques de son évolution et de ses avancées. Il dépend du génie linguistique au sens large de pouvoir créer les métalangues propres à maîtriser les nombreux objets sémantiques ou textuels d'une grande bibliothèque ou d'un grand réseau de communication électronique (au sens de rassemblement et de coopération de grandes masses de documents numérisés).

L'informatique permet de capitaliser les recherches diverses qui sont menées. Des montages diachroniques complexes peuvent être envisagés, plus proches des montages cinématographiques en *feed back* et projection par exemple, que de suites linéaires séquentielles. Les simplifications et l'agrégation des noeuds des principaux faisceaux en bibliothèques de liens au sein de connexions hypertextuelles, permettent d'envisager de telles avancées.

Il ne s'agit donc pas de rester sur l'idée que l'informatique apporterait une simple alternative formelle et de reproduction artificielle à l'imprimé. Ce qui est en jeu, c'est la rapidité, l'efficacité même du cycle de l'information entre son émission, sa publication et d'autre part sa réception, son appropriation par un public concerné. Vannevar Bush insistait déjà sur la lenteur de ce cycle dans le cas de compréhension de découvertes scientifiques, par exemple avec les lois de Mendel en génétique, "comprises" cinquante ans seulement après leur publication ! Ceci est encore plus évident dès lors que l'on considère la complexité des offres d'information sur les grands réseaux de type Internet où, faute d'une logistique de navigation, d'orientation, les informations se banalisent, perdent de leur relief "réticulaire", par impossibilité de les saisir ou les traiter dans le temps réel de l'interactivité.

Dans un premier temps, entre les entrées par le catalogue et le chargement de documents provenant de la collection numérisée, toute une architecture d'aide à la décision peut être précisée : banques de données bibliographiques, collection en mode texte, moteur de recherche, sur des extraits enrichis des notices (table des matières, index, listes bibliographiques, etc.).

Mais à l'avenir, le devenir du livre et de l'imprimé, donc des bibliothèques, se jouera en fonction des possibilités réelles que l'informatique saura développer, non pas seulement pour cataloguer, décrire, reproduire, mais aussi pour lire avec l'assistance de fonctions nouvelles, pour représenter les possibilités de recherche et d'orientation, pour cartographier les nouvelles configurations du savoir.

On mesure l'importance pédagogique de telles transformations.

Les bibliothèques électroniques de demain, dans une vision patrimoniale dynamique, auront donc en charge ce travail de mise en évidence des communautés virtuelles qui souligneront l'importance culturelle de tel ou tel débat, de telle ou telle oeuvre créatrice.

Ainsi une des missions patrimoniales à venir serait de donner l'étalon diachronique des publications, d'en faire une sorte de curseur fondamental (pas nécessairement linéaire), mettant en scène les réarrangements disciplinaires que subissent et provoquent les avancées (et les disparitions !) des connaissances. Imaginons un instant de tels parcours, voyant apparaître tel pan des découvertes géographiques, zoologiques ou ethnographiques, voyant disparaître telle référence institutionnelle, d'ordre monastique, d'école alchimique, ou de jurisprudence ! Nous voyons là se réorganiser les *épistémè*, aux marges des connaissances déjà acquises... L'utopie est ici d'une exigence sans borne au regard de la pauvre informatique dont disposent les appareils documentaires actuels !

Une nouvelle question économique pour la chaîne éditoriale

Cette dérégulation au plus près du texte en tant qu'ensemble publié touche toute la chaîne éditoriale. En conclusion, la question est posée à l'ensemble des acteurs de cette chaîne de trouver des solutions techniques, économiques, financières, en accord avec cette dimension de création virtuelle.

Là encore, Ted Nelson fut précurseur. Tout document est privé ou public, car tout document a un auteur, et cette propriété se met en oeuvre par des autorisations ou des impossibilités d'accès, mais aussi par un *copyright* qui accompagne chaque document signé s'il fait l'objet d'une publication, dans une dynamique sans frontière des circulations et des connexions. De ce fait, tout document peut ensuite être repris, approprié, modifié, Nelson ne mettant aucune limite à ce développement de la culture dès lors que le *copyright* est respecté. Le statut légal de l'information électronique ne fait que découler de cette approche culturelle et économique. Pour l'auteur de *Xanadu*, il n'y a pas de système de documents électroniques sans règles de redevance financière et sans établissement des coûts d'utilisation.

La sphère de la communauté devient coextensive à cette nouvelle diffusion du savoir, élargie, approfondie. Les fonds virtuels, construits par interconnexion, sont le fait des universités, des centres de recherche, des maisons d'édition, qui programment et proposent sur les réseaux des agencements inédits, des "émissions", (pour emprunter le terme de l'unité documentaire en cours sur les chaînes TV), revisitant et réutilisant à l'infini des sources patrimoniales. Peut-on visiter votre bibliothèque virtuelle ? sera une question souvent posée sur les réseaux, pour accéder au fonds de référence de telle démarche scientifique, de telle entreprise, de tel projet d'écriture ou de tel cas de jurisprudence.

On l'a souvent dit, l'électronique fait passer le texte d'une économie de l'objet produit, stocké (le livre) à une économie du flux d'informations. On l'a aussi relevé : la première suppose un rapport d'échange où l'achat marchand est préalable, *a priori*, et signe un droit de propriété quasiment sans limite sur la jouissance de l'objet (relire sans cesse le livre, le prêter...). La seconde suppose et nécessite une économie du paiement à la consultation en flux et une gestion du partage et du reversement des redevances entre auteur et éditeur, opérateur du service, transporteur des communications. Mais, aussi bien en amont, où l'auteur risque de se diluer dans un auteur collectif indifférencié, qu'en aval, où la consommation en ligne démultiplie à l'infini les accès, cette économie réticulaire de la demande doit reposer sur de nouveaux découpages des espaces de gratuité publique, pour promouvoir la connaissance, l'éducation et la recherche, mais aussi sur de nouveaux espaces de profits liés à ces nouveaux métiers.

Ces nouvelles frontières ne peuvent se dessiner que par de nouvelles règles contractuelles et une redistribution des responsabilités de tous les acteurs en jeu : pour l'heure, souvent crispés sur les méfaits de la technique (et ce depuis la reproduction en photocopie), peu capables de discerner et de mettre en oeuvre les marchés potentiels et les bienfaits de l'édition électronique, les fabricants du livre, les pouvoirs publics, les conservateurs de patrimoine, campent, forts de leurs intérêts légitimes, sur le terrain connu de l'économie de l'objet, sur un terrain national et linguistique bien balisé. Mais qui ne voit que ceux-ci sont circonscrits, circonvenus par les réseaux, internationaux, multilingues ? Michel Serres a raison, la seule économie qui vaudra sera celle qui collera au plus près de ce fonctionnement en flux des documents.

Bruit et sens

Bruit et sens : au fond, que veut dire publier, quand l'élaboration d'un texte électronique, totalement composé, peut être le fait de tout un chacun, d'une communauté en devenir ? Il faut bien que de nouvelles pratiques éditoriales trient, mettent en valeur, de façon plus sélective, en raison même des débordements informationnels. Relation de l'inédit au patrimoine, réécrire sans fin de façon redondante ou rareté de l'intérêt ?

Le statut légal de l'information électronique sanctionnera surtout cette nouvelle économie de flux, de coopération, de valorisation, sur la base d'un accroissement sans égal des consommations d'information, à très faible coût unitaire, et indépendamment des contraintes économiques liées à l'imprimé physique.

Pierre Lévy a raison d'écrire : “ La bibliothèque encyclopédique chasse le Livre. Et la bibliothèque s'étend, déborde, tente de se repérer elle-même par ses fichiers et ses index. Les citations et références dessinent un réseau des lectures, des circulations dans le savoir qui rongent les anciennes frontières. Bientôt, les journaux scientifiques se multiplient, déversent des marées d'articles, qui alimentent à leur tour des banques de données innombrables. Comment classer, organiser l'information scientifique et technique, mais aussi bien les images des musées, des archives audiovisuelles et le flot montant de l'information en général ? Les efforts mala-droits pour faire encore servir à cet effet les corps raidis des vieilles disciplines ne peuvent aboutir qu'à l'échec, quand le savoir ne vit qu'aux marges mouvantes, aux carrefours dans les interférences, quand tout devient affaire d'import-export ”²⁰.

Si ces techniques, cet art de la navigation et de la valorisation éditoriale composant ces nouvelles bibliothèques virtuelles, ne parviennent pas à s'imposer, de nouvelles formes d'oubli, d'anéantissement seront liées à l'écrit électronique, par excès, déperdition de traces, banalisation des références, nivellement des *data*, perte du sens pour sombrer dans le bruit indifférencié.

C'est aussi qu'une communauté savante, éducative, éditoriale aura échoué dans l'utilisation informatique. Plus que jamais, les synthèses éditoriales, sur les réseaux, devront sauver de l'oubli en orientant, spécifiant, accompagnant les recherches documentaires et le traitement de masses d'informations sans égales.

“ Quand le stock s'identifie au flux, les grandes concentrations se dispersent en singularités ”²¹. La dialectique du local et du global se redéfinit par le biais des connexions du réseau, seule dimension virtuelle gérant les proximités d'intérêt intellectuel, au-delà des séparations de l'espace et du temps.

²⁰ Pierre LEVY, Ibid., p. 201.

²¹ Michel SERRES, Ibid., p. 153.

Les grandes bibliothèques à l'avenir, dans leurs stratégies, sont vouées, par la grâce de la numérisation à une économie de flux plus que de stock, de communication plus que de conservation, de réseau plus que de volonté patrimoniale.

C'est à ce prix que la bibliothèque électronique sauve du bruit informatique, du risque d'anéantissement de la mémoire, et participe d'un encyclopédisme renaissant.

Janvier 1995