

Espace. Temps. Information

V. L'hypertexte comme système comportemental

LAB[au], Manuel Abendroth, Jérôme Decock

lab-au@lab-au.com

Salvino A. SALVAGGIO, Ph.D.

salvino@salvaggio.net

D'un point de vue purement technologique, l'hypertexte est un système de transmission et d'organisation des informations basé sur des processus de computation des données (code) et sur la mise en place d'interconnexions (structure/réseau). Il s'agit d'un mode de communication dépendant d'une technologie qui délivre non seulement un message (contenu) mais également, à travers son code et sa structure, un sens spécifique. Ce sens spécifique, contenu dans la méta-construction de l'hypertexte, le définit comme un méta-langage, un mixte de technologie, de structure et de contenu, qui influence profondément notre manière de communiquer, de comprendre et de percevoir.

L'Hypertexte s'organise sur base d'interconnexions (hyperliens) et de processus de computation entre des unités d'information. Si la technologie en elle-même est porteuse d'une signification qui conditionne notre manière de formuler et de comprendre un message¹, l'hypertexte comme medium va plus loin en représentant par sa structure d'hyperliens des relations mentales et cognitives, une reproduction de schéma de pensée qui fonctionne par association (hyperlien). L'hyperlien amène à l'introduction productive de l'utilisateur dans le système d'interconnexion (l'interactivité) et de l'hypertexte comme un système d'enaction partant de la mise en relation de pensée et action. L'hypertexte constitue ainsi un langage d'actions productives basé sur des processus cognitifs et mentaux qui en "enactant" une signification peut être rapproché des modèles de fonctionnement neuronal. Il apparaît donc comme un système comportemental.

"As we may think"
Vanevar Bush, 1945

A l'origine des sciences de l'information, le MEMEX de Vanevar Bush fut une des premières études menées sur un objet technologique dont le but était d'accroître nos capacités mentales et cognitives. Conçu comme l'extension de notre mémoire, le MEMEX mettait en place quelques-uns des principes de l'hypertexte, en particulier en proposant un système d'archivage basé sur des associations (hyperliens), un processus non plus exclusivement relatif au contenu mais aussi à la qualification/structuration qu'on lui applique. Ce processus de traitement ou d'indexation de l'information impliquait dès lors --par la logique computationniste² et les conditions d'interfaçage homme/machine-- la naissance et le développement

¹ Voir article Inside Internet n°34, juillet 2000

² Logique mise en place pour expliquer la perception par nos sens, les traitements multiples de l'information, la cognition,... Au début des années 50 au MIT, Norbert Wiener, le père de la cybernétique, fut l'une des personnes qui travailla sur cette logique, à la base des sciences cognitives et de l'information.

conjoint des sciences cognitives et des sciences informatiques, surtout si l'on considère que le but d'un système comme le MEMEX était de sauvegarder un schéma de pensée, une structuration temporaire de l'information, comme le fait notre mémoire.

Le MEMEX suit ainsi la logique computationniste déterminant la cognition de notre cerveau, comme un traitement d'information (computation), un système qui effectue sur des symboles (sémantique) des opérations selon des règles (syntaxe).

En effet, dans notre mémoire, l'information est stockée sous forme d'unité de sens, des représentations construites, des blocs identifiables qui ne sont néanmoins jamais totalement indépendants d'un contexte et d'une série d'associations que nous leur appliquons de manière consciente ou inconsciente. Ce processus est devenu dans le MEMEX une méthode d'indexation, non plus relative au contenu lui-même (ce que nous retenons) mais plutôt au comment et pourquoi nous retenons ce contenu (processus cognitifs). Dans le cas du MEMEX, comme dans celui de l'hypertexte, les associations ou les hyperliens fournissent un " contexte " au contenu spécifique de l'unité d'information, une série d'informations qui lui sont relatives. Ces hyperliens constituent donc une information en eux-mêmes mais une information qui apporte un sens complémentaire au sens propre d'un message, un sens basé sur des associations (structure). Ce sens contenu dans l'articulation particulière de fragments d'information exprime certes une logique structurelle mais surtout un individu. La logique d'associations libres définie par l'utilisateur représente la structuration d'une pensée qui lui est propre et qui le qualifie. Au contraire des systèmes d'archivage et d'indexation précédents qui se sont tous vus plus ou moins universels, le MEMEX fut conçu comme un objet personnel dont le fonctionnement se basait sur les processus cognitifs constituant le fonctionnement " de base " du cerveau. Dans leur dimension " archivage ", le MEMEX ou l'hypertexte peuvent ainsi se définir comme des extensions du cerveau humain, une structure décrite par les processus qui catégorisent et lient les idées par une indexation associative du savoir.

En revanche, et contrairement au MEMEX, l'hypertexte n'est pas un système personnel/individuel ou un système d'indexation manipulé par un utilisateur et destiné à la seule consultation de celui-ci, mais une modalité de communication donc un système interpersonnel. Elevé au statut d'un langage d'échange, un protocole de transfert généralisé, l'hypertexte se comprend comme un système de communication introduisant l'utilisateur comme partie productive et corrélative du système d'interrelation. La construction d'un hypertexte implique donc la mise en place consciente du contenu et des liens par un " auteur " avec l'intention d'être perçu par d'autres personnes. En pratique, ces relations entre " informations " (textes par exemple) se trouvaient déjà dans le texte imprimé et ont été dénommées sous leurs multiples manifestations par les concepts d'*intertextualité*³ ou de *métatexte*⁴, mais l'hypertexte, et la technologie qu'il emploie, combine ces relations à une expérience qui unit le lecteur au médium (interactivité). Si cette association d'informations (la

³ La présence effective d'un texte dans un autre, du plus explicite ou moins explicite, la citation, le plagiat, l'allusion (qui pour être effective suppose que le texte ne peut se comprendre que si le lecteur est en mesure de décoder l'information et qui si elle n'est pas perçue le message est irrecevable). Voir " Sêméiôtikè " de Julia Kristeva, Seuil, 1969.

⁴ La métatextualité est la relation qui unit un texte à un autre texte dont il parle, sans nécessairement le citer, voire sans le nommer. Une critique de Roman est par exemple un métatexte . Voir " *Palimpsestes, la littérature au second degré* ", de Gérard Genette, Seuil, 1982.

structure-contexte) produit un sens, c'est essentiellement parce qu'elle est soumise à l'interprétation et à la cognition de chacun. L'interprétation comme la cognition se définissent comme le rapport mutuel que l'on peut établir entre une source d'information (structure et contenu) et l'association (interprétation/représentation) ou la reconnaissance (cognition) du sens qu'elle porte. Ils sont des processus d'extraction du savoir qui font intégralement appel à la participation active de l'utilisateur et à son implication dans l'élaboration d'une communication, un rapport défini dans l'hypertexte par l'interactivité et le rapport d'*enaction* qu'il met en place. Par conséquent, la mise en place de liens dans un hypertexte permet non seulement aux auteurs mais également aux lecteurs de créer des associations, entre différentes unités d'information ou entre différentes idées relatives, qui ont un sens pour eux. L'hypertexte n'est donc pas un simple système d'archivage, même si cet archivage incorpore en soi la représentation de schémas de pensée; il est un système comportemental incluant des processus cognitifs. En d'autres termes, l'utilisateur est totalement impliqué dans une expérience (perception) qui à travers l'interactivité fait appel à ses capacités cognitives et à son interprétation (association / choix), un processus qui peut révéler ses intérêts et surtout son comportement.

Par conséquent dans le domaine des nouveaux media, il importe de comprendre la relation qui s'établit entre d'une part la perception (l'utilisation des sens), la reconnaissance ou la compréhension et la représentation (l'extraction du sens), et d'autre part l'action qui en résulte (la production du sens).

En soi le parcours d'un hypertexte ne produit pas que de l'information, il produit, au travers de processus cognitifs, un sens dont l'extraction dépend de la technique de computation utilisée ou du but recherché. Ainsi, parmi les techniques qui suivent et enregistrent les parcours des utilisateurs sur Internet figure le tracking et le profiling.

Le tracking opère comme un dispositif qui collecte des informations sur des utilisateurs d'une manière " statique ", de la configuration matérielle de leurs ordinateurs jusqu'aux informations personnelles, ou d'une manière " dynamique " en suivant à la trace leurs évolutions dans le réseau. Le but du tracking est essentiellement d'établir des groupes cibles, d'analyser statistiquement les informations collectées pour en déduire des classes sociales, des groupes d'intérêt et aussi le comportement des utilisateurs, c'est le mapping des utilisateurs à l'échelle du groupe. Cette analyse est souvent utilisée pour améliorer la conception et l'ergonomie d'un site ainsi que pour établir un plan marketing à des fins de publicité ciblée. Ce procédé s'effectue typiquement sans que l'utilisateur en ait conscience.

Les techniques mises en œuvre dans le profiling sont similaires mais les buts et surtout l'utilisation faite des informations changent. Plus que le tracking, le profiling est une analyse du parcours de l'utilisateur dans un hypertexte ou dans un réseau. Ce parcours à travers l'information structurée produit un sens spécifique. L'importance et la validité du profiling dépendent donc de l'indexation des contenus, de la mise en place de liens ou de systèmes graphiques dont l'impact cognitif et l'interprétation sont connus ou reconnus. Le profiling peut certes s'établir sur les mêmes bases statistiques que le tracking mais aussi par une analyse sémantique des interactions par rapport à l'indexation (sémantique et cognitive) des contenus. Le profiling consiste donc à extraire le sens produit par les interactions de l'utilisateur dans le but de comprendre ses comportements. Il s'agit du mapping des utilisateurs

à l'échelle individuelle, de la construction d'un profil sociologique jusqu'à un profil psychologique, et du reformatage de l'information en fonction de ce profil. En effet le profiling inclut des éléments proactifs comme l'affichage " personnalisé " de certaines informations prioritaires par rapport à son intérêt plutôt que d'autres informations généralistes.

Le profiling contient donc des méthodes qui, en qualifiant l'utilisateur par ses interactions, pourraient aboutir à la constitution d'une véritable identité " virtuelle ". Mais de telles démarches ne peuvent réussir que si le processus de profiling se passe en toute transparence et que l'édition des informations contenues dans le profil est permise. En gagnant en accessibilité ou en tirant profit d'agents intelligents mis à sa disposition, l'utilisateur pourra valider son profil par une utilisation et un intérêt commun avec les producteurs de contenu.

L'hypertexte comme système comportemental basé sur des processus mentaux et cognitifs (neuronaux) affectent autant le corps social (langage) que l'individu. Comme langage il reconfigure non seulement des " schémas de pensée " (sens - sémantique) mais aussi nos sens (perception). L'hypertexte, précisément parce qu'il assemble des processus de computation et de communication dans un seul système cohérent, constitue un modèle d'analyse, d'organisation et d'interaction spatio-temporelle de l'information, démontrant les implications des nouvelles technologies dans la société de l'information et des réseaux.

Privacy, the consumer information organisation

<http://www.privacy.net/>

Les site de Privacy démontre par l'exemple quelles sont les informations qui peuvent être collectées par le simple fait d'accéder à un site ou d'activer un lien. Rendues visibles, la valeur commerciale et l'accessibilité de ces informations ne fait plus de doute quant à leur utilisation ce site fait le tour des exploitations et des méthodes et pratiques déjà établies. Avoir conscience que le tracking et le profiling font parties des méthodes courantes sur Internet n'en révèle pas que les aspects déontologiquement discutables mais aussi les moyens de les détourner et le potentiel qu'elles présentent dans l'accès et la personnalisation de l'information.

Memex, Vannevar Bush

<http://www.dynamicdiagrams.com/design/memex/memex.htm>

The Brain –personal brain (voir article Inside Internet n° 34, juillet 2000)

<http://www.thebrain.com>

" Juillet 1995 fut marqué par le cinquantième anniversaire du Memex, le modèle/prototype des ordinateurs personnels d'aujourd'hui, décrit par Vannevar Bush dans " As We May Think ", son essai de 1945. Dynamic Diagrams a produit une animation qui illustre les idées de Bush, en suivant la structure de l'essai. " Dans cette présentation d'un objet qui n'a jamais existé, le Memex, Dynamic Diagrams, une structure fondée par un graphiste et un écrivain, soulève, à travers une simulation quel était son usage pratique et ses potentiels. Si sous certains aspects, le Memex peut sembler dépassé dans ces options technologiques, il apporte une vision très révélatrice des principes d'indexation hypertextuelle et de computation qui

furent et qui restent ceux de l'ordinateur personnel. Le *Personal Brain* en proposant une indexation associative de documents propose une approche similaire au *Memex*, adaptée à la technologie actuelle.

DMOZ – opendirectory project – the Internet Brain

<http://www.dmoz.org>

WebBrain

<http://www.webbrain.com>

Sous le slogan “ humans do it better ”, ce site prône une indexation ‘intelligente’ du contenu d’Internet, que la technologie actuelle es , selon eux, incapable d’effectuer. Des technologies comme celles présentées ici et l’avancée des sciences cognitives conduiront peut-être un jour de contredire cette vision. Néanmoins le système d’indexation proposé par Dmoz est un système qui tire son efficacité d’un consensus et d’une intelligence collective ; il est en permanence enrichi et actualisé à travers le réseau par des centaines d’utilisateurs (web editors). En s’associant récemment avec Brain pour former le WebBrain, Dmoz rend visible les multiples liens qui existent entre les catégories et permet une navigation multilinéaire basée sur une indexation ‘interactive’ propre à la société des réseaux.

Autonomy Kenjin

<http://www.autonomy.com/>

<http://www.kenjin.com>

La technologie mise au point par *Autonomy* fournit une infrastructure automatisée pour le profiling des utilisateurs, la catégorisation, l’association et la personnalisation de l’information non structurée. Ce moteur d’indexation est basé sur des algorithmes et les probabilités. En analysant une information, son contexte et son occurrence, l’ordinateur devient capable d’estimer (probabilité) qu’il y a X% que le document en question soit à propos d’un tel ou d’un tel sujet. La définition de mots-clefs comme l’indexation sémantique et associative de tout document deviennent dès lors automatiques. Le moteur de recherche *Kenjin* est un programme qui en analysant le contenu de tout document consulté (localement et sur internet) construit une base de données des intérêts ou un profil d’une personne. Plus grande est la base de données, plus le programme devient apte à comprendre l’utilisateur dans ces recherches et même à lui suggérer des contenus.

Conversation Map, Warren Sack

<http://www.media.mit.edu/~wsack/CM/index.html>

Cette cartographie d’une conversation dans un *chat-space* indexe les échanges selon 4 critères ; 1. *Social mapping, Social network* - les relations entre les personnes, 2. *Thématique/sujet* – l’importance des thèmes dans la conversation 3. *Sémantique* – les relations entre les thèmes abordés, 4. *Chronologique*. Ces 4 visualisations interactives présentent une recherche orientée sur l’optimisation de la communication dans des conversations “ many to many ” tel que les proposent actuellement les forums de discussion. En indexant les échanges par rapport au temps, ce programme permet de déterminer la position relative d’un individu ou d’un groupe d’individu ; en indexant les thèmes et leur relation sémantique, d’identifier leurs intérêts et leurs associations (sujets/processus de conversation).

A travers cette cartographie et les systèmes de consultation qui lui sont associés, c'est en fait un véritable hypertexte qui se génère, en combinant indexation et communication, ce système révèle le statut de " langage " de l'hypertexte.