



# Le Web ou l'utopie d'un espace documentaire

Dominique Boullier, Franck Ghitalla

► **To cite this version:**

Dominique Boullier, Franck Ghitalla. Le Web ou l'utopie d'un espace documentaire. Revue I3 - Information Interaction Intelligence, Cepaduès, 2004, 4 (1). <sic\_00001022>

**HAL Id: sic\_00001022**

**[https://archivesic.ccsd.cnrs.fr/sic\\_00001022](https://archivesic.ccsd.cnrs.fr/sic_00001022)**

Submitted on 5 Jul 2004

**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.

# Le Web ou l'utopie d'un espace documentaire

Dominique Boullier, Franck Ghitalla

COSTECH, Université de Technologie de Compiègne,  
Centre Pierre Guillaumat – Royallieu  
rue du docteur Schweitzer, B.P. 60 319 Compiègne CEDEX.

[dominique.boullier@utc.fr](mailto:dominique.boullier@utc.fr)

[franck.ghitalla@utc.fr](mailto:franck.ghitalla@utc.fr)

## *Résumé*

*L'exploitation raisonnée du web représente un défi pour nombre d'utilisateurs tant s'y côtoient les modèles de documents mais aussi s'y exercent les difficultés techniques. C'est toute l'activité critique, et plus particulièrement la « lecture », qui se trouve alors remise en question étant donné l'incertitude native de cet espace de navigation. C'est ce qu'aura montré une enquête conduite sur près de deux ans en suivant dans leur activité une quarantaine d'utilisateurs.*

**Mots-clés :** *web, document, observation, usages, stratégies de navigation*

## *Abstract*

*The exploitation of the Web for reasoning tasks represents a challenge for a great number of users. Their activity is shaped by technical difficulties and by the necessity to control a lot of types of documents. Thus, in this kind of undefined navigation space what we call « reading » becomes a quiet complex activity to describe. It is one of the most important conclusion of a study of two years.*

**Keywords :** *web, document, activity, study, social navigation, reading tasks.*

# 1 INTRODUCTION

Durant près de deux ans, nous avons observé une quarantaine d'internautes aux profils divers dans leurs usages quotidiens du web, aussi bien sur station publique (Bibliothèque Publique d'Information, Cyberbase de la Cité des Sciences) qu'à leur domicile [7]. L'objectif initial de l'enquête était de relever suffisamment d'indices pour comprendre comment se renouvelait aujourd'hui sous nos yeux l'activité de lecture, en particulier sous l'influence des supports numériques et de leur circulation sur les réseaux. Au terme de l'enquête, il apparaît que ce n'est pas seulement la lecture qui est concernée mais l'ensemble des cadres qui la rendent possible. Qu'on l'envisage sous l'angle restreint du déchiffrement des signes ou de façon plus large comme activité interprétative, la lecture suppose en effet une aptitude technique à maîtriser les formats d'un environnement matériel et, plus largement, la possibilité de l'inscrire dans une série de conventions qui en règlent les formes et les occasions. Ce sont précisément ces cadres techniques et sociaux incertains sur le web qui ont massivement fait difficulté pour nos usagers, hypothéquant sérieusement leur activité critique. Malgré les difficultés, et par delà la diversité des pratiques et des stratégies pour s'y retrouver sur le web, ils ont effectivement « lu » à certains moments mais *seulement* à l'intérieur de cette série organisée de cadres que constitue tout document.

La question de la lecture sur le web a donc emporté avec elle celle du document, de ses modèles hérités de l'imprimé, de ceux qui émergent avec les supports numériques ou encore de ceux qu'il a fallu construire *ad hoc* pour des usagers condamnés le plus souvent à les faire tous alterner, quel que soit leur niveau supposé de compétence en matière de navigation. Salutairement, les difficultés que nous avons rencontrées pour décrire leur comportement et leurs différents modèles interprétatifs du « document web » nous ont obligé à interroger la pertinence des termes et des concepts que nous utilisons spontanément, ceux avec lesquels se conçoivent les différents aspects d'une architecture documentaire : nos observés ont effectivement « lu » mais quel type d'objets exactement : des « bouts » de textes, une « page » HTML, une mosaïque d'objets sémiotiquement disparates, la surface d'un écran, une suite de fenêtres ouvertes successivement, des « événements interactifs »? Dans un espace aussi peu structuré que le web, des modèles de documents se côtoient ou se succèdent pour l'utilisateur, eux-même formant des « îlots de cohérence » au milieu des communications les plus diverses. L'observation des pratiques de lecture sur le web s'articule ainsi logiquement à la description des « styles cognitifs » des usagers et des stratégies qu'ils

adoptent pour gérer la diversité mais aussi l'incertitude de l'espace qu'ils parcourent.

## 2 LES FORMATS DU DOCUMENT NUMERIQUE

« C'est facile, dit Alain, quand on clique on a tout ce qu'on veut ». Et il n'y a plus qu'à « lire le document » à écouter certains des observés. En réalité leur activité a été le plus souvent émaillée d'incidents et de difficultés qui ont hypothéqué l'accès à une information recherchée ou les mécanismes mêmes de l'interprétation de ce qu'ils avaient sous les yeux. Ces données ont été si récurrentes au cours des observations qu'il a bien fallu reconnaître que la « lecture » d'un document passait aussi par la maîtrise d'un espace techniquement organisé. La rencontre avec un support numérique aux fonctionnalités inédites montre exemplairement comment se combinent interprétation du *sens* et manipulation du *format matériel* de l'espace documentaire. Le terme de « format » s'est trouvé d'autant plus justifié dans notre analyse qu'il renvoie non seulement à la dimension de l'activité chez les usagers mais aussi à la façon dont au cours de l'histoire cadres sémiotique et technique se sont conditionnées mutuellement et dont nous avons hérité de formes canoniques (le roman au format de poche, le slogan et l'affiche, titraisons et presse). Nos observations renvoient massivement à ce que R. T. Pédaque souligne comme « la manifestation la plus évidente du changement » à l'ère du numérique, c'est-à-dire « la perte de la stabilité du document comme objet matériel » [1]. On souligne ainsi dans son document que la notion de support n'a jamais été aussi « problématique » tant se voient modifiés les cadres qui nous font percevoir un « document » comme une forme organisée et reconnaissable. Notre enquête le confirme et l'une des questions majeures qu'elle aura suscitées est bien celle des modes de clôture du document numérique, de sa morphologie et de ses formes temporelles. Déjà entamée ailleurs [9], la question revient avec d'autant plus d'acuité théorique qu'elle conditionne, selon la perspective que nous avons adoptée, la nature même de l'activité critique et de ses formes logiques [11][18].

### 2.1 La couche interactive

La maîtrise des diverses formes d'interactivité nous a semblé être la

condition *sine-qua-non* de toute activité interprétative, en soulignant que l'on entend par « interactivité » autant les manipulations techniques qui lui sont associées que la dimension sociale de l'interaction et de ses conventions naissantes. Les usagers ont ainsi passé beaucoup de temps à repérer les « zones réactives » ou à essayer de comprendre les principes de leur distribution dans l'espace. Comme le dit Catherine : « Et il faut rechercher, attendre, voir où l'écran devient sensible. Normalement, c'est comme sur les cédéroms, il y a un *point sensible* ». Nous avons souvent eu l'impression que seul le balayage de la surface de l'écran à l'aide de la souris guidait l'activité. Parfois par plaisir, ou simple jeu. Le plus souvent par nécessité pour l'hypothèque que l'interactivité fait peser sur la construction du sens et la compréhension générale des cadres du document sur support numérique. Cette dimension ne s'arrête d'ailleurs pas au seul lien hypertexte et aux surprises sur lesquelles il ouvre parfois : des *roll-over* aux fenêtres de saisie de texte en passant par l'activation de scripts, c'est l'ensemble de la géographie du document qui se trouve tissée pour nos usagers « d'aspérités réactives ». En retour, les usagers ont souvent l'impression de « calculer l'information », ou du moins que « ça calcule » pour eux. En particulier, ils ont accordé une grande importance au rôle des requêtes et à leur forme linguistique. Dans la géographie de l'interactivité, les zones de saisie, par exemple, semblent avoir encore plus d'importance que le repérage des ancrs hypertextes. C'est notamment le cas pour ces « objets » typiquement numériques que sont la barre d'adressage du navigateur, les écrans de requête des moteurs de recherches et ceux qui sont directement intégrés aux pages web à titre de fonctionnalité propre d'un site. Peut-on encore parler dans ce cas de « document » ou s'agit-il simplement de fonctionnalités techniques liées au dispositif de navigation? La dimension de l'action sur les signes et d'une possible *fabrication automatique du sens* fait pourtant partie d'une conception spontanée du document numérique comme une *machine à fabriquer de l'information*, profilée pour leurs propres besoins. Ainsi pour Maurice « ça tourne bien » quand il valide une zone de saisie ou pour Didier pour qui le moteur interne à un site « continue de *mettre en boîte* les informations ». Hormis la phase préliminaire de recherche dans un moteur, la distinction n'existe plus pour eux entre outils d'interrogation et d'indexation et contenu affichable comme résultat de leur manipulation.

Ces aspects de l'activité spécifiques de l'exploitation technique du document numérique posent la question de savoir si des « modèles rhétoriques » de l'interactivité ont pu être esquissés au cours de l'enquête, à l'image de cadres canoniques naissants. Nous avons relevé une série d'indices significatifs qui semblent guider nos usagers comme régler des solutions de conception réparables. En particulier, l'utilisation et

l'exploitation de l'hypertexte et de certaines fonctionnalités interactives (comme les *roll-over* et la concentration des zones interactives en haut ou à gauche des pages HTML) semble s'accompagner de l'apparition d'un type spécifique d'informations destinées à les signaler comme tels (du type « cliquez-ici »), comme si les signes ou le document devaient intégrer leur propre *mode d'emploi*. Ivana, l'une des observées, se sent « bloquée » devant « l'absence d'un logo-lien » à la fin de la page qu'elle est en train de lire, regrettant de ne pouvoir « trouver en fin de page » ce qu'elle juge indispensable. Des modèles implicites d'organisation de l'espace signifiant semblent être en train de se construire, par expérience accumulée et par référence à des sites qu'ils connaissent bien. A ce titre, un site est vite jugé « mal fait » s'il diffère trop du modèle attendu. Ces attentes implicites des usagers en termes d'adoption de solutions spécifiques de composition sémiotique montrent que c'est aussi l'accès à l'information, et au delà au sens, qui se joue. A ce titre, on aura noté que l'un des effets interprétatifs majeurs de l'interactivité chez les usagers repose sur la généralisation de la recherche d'indices, ouvrant le document de façon massive à la dimension de l'association au moins autant qu'à la discursivité de l'argumentation. En d'autres mots, sur un support numérique, tout *signe* (image, mots, lettres, symboles, zones colorée) se voit doté par l'utilisateur d'une valeur d'*indice* d'une possible association avec un autre contenu documentaire. Ainsi des usagers ont-ils cherché à activer des zones inertes tandis que d'autres ont pris soin de balayer la surface de l'écran avant toute entreprise de déchiffrement. Et les effets de ce principe sont loin d'être négligeables : on aura ainsi remarqué qu'avec l'interactivité les mots ont parfois plus d'importance que les phrases, autrement dit le réseau discursif dans lequel ils sont pris. Si l'on veut bien admettre qu'il existe une forme de lecture de « balayage » ou de « scannage » où l'on cherche sémantiquement à repérer des « mots-clefs », on comprendra que dans ce cas l'interactivité renforce le principe de la *granularisation* de certains mots qui n'ont alors plus le même statut que les autres. Si l'on a beaucoup insisté sur la modélisation des tâches cognitives impliquées dans la « recherche d'information », il serait aussi pertinent de s'interroger sur les effets de la généralisation d'une sorte de *paradigme indiciaire* [10] qui fait de l'écriture comme un ensemble « de signes assoupis qu'on doit réveiller pour qu'ils se mettent de nouveau à parler »<sup>1</sup>. Pour nos usagers l'indice d'interactivité est toujours la promesse d'un « accès direct à l'information » comme dirait Jeff.

---

<sup>1</sup> Selon la phrase de M. Foucault pour qualifier la « quête des similitudes » du Quichotte. *Les mots et les choses*, Gallimard, 1966, p.61.

## 2.2 Modes de clôture et fenêtrage

Au delà des mécanismes de l'interactivité, c'est toute la question de la géographie du document web qui se trouve posée. Si nous avons l'habitude de parler d'une « page » web par référence à l'imprimé ou à la structure du script HTML, il n'est pas sûr que ce soit là l'unité de mesure qu'ont adoptée les usagers pour « lire », parcourir ou même naviguer. On pourrait alors voir en l'écran et en ses possibilités d'affichage l'espace reconnu à l'intérieur duquel se déroule l'activité, s'effectue la lecture, se construit le document. Or ni l'un ni l'autre ne constituent cette unité de mesure qui donne à la géographie de l'activité ses points de focalisation et son rythme mais bien plutôt la répartition de l'espace en fenêtres, actives successivement, s'effaçant les unes les autres ou se superposant. L'enquête a montré que c'est d'abord des stratégies de gestion du fenêtrage que dépendait le déploiement de l'activité interprétative. L'observation de l'activité dédiée à la gestion du *windowing* fournit en effet un certain nombre d'indices sur la morphologie particulière du document web. E. Souchier et Y. Jeanneret ont déjà repéré l'organisation hiérarchique des différents types de fenêtres et qu'ils nomment « les cadres de l'écrit d'écran » [13][19]. Il s'agit d'une théorie de la mise en abîme des cadres constitutifs d'une ingénierie textuelle où s'enchaînent « en profondeur » cadres de l'écran, du « bureau » (ou du système d'exploitation), du logiciel (par exemple un traitement de texte) et pour finir de cet espace « scriptible » constitué par la fenêtre active. Ces cadres d'une « énonciation éditoriale » en couches déterminent ainsi différents niveaux de l'activité pour l'utilisateur auxquels il faut adjoindre une analyse de leur succession temporelle pour comprendre la « forme » d'un parcours de navigation, ses patterns spatiaux et les stratégies cognitives qui les guident.

Un second principe fort semble gouverner l'activité et la géographie interne du document : il s'agit de la distribution des signes en différentes zones à l'intérieur des fenêtres. Cette composition en tableau conduit en effet les usagers à les considérer séparément, voire à leur accorder des degrés de hiérarchie. Les pages web sont en effet souvent regardées comme des tableaux dont les cases seraient relativement autonomes les unes par rapport aux autres. L'utilisation massive de nos jours des tags HTML de type *TABLE* et de tous ses attributs montre à l'évidence que l'on recherche à renforcer une stabilité géographique au niveau de l'affichage qui assure aux éléments constitutifs de la page la préservation de leurs rapports plastiques mutuels. Mais la distribution « en tableaux » dont il s'agit concerne le plus souvent seulement l'organisation

géographique du document ; on y vise surtout à distribuer l'autonomie de ses composants internes et très rarement une méta-organisation de relations logiques croisées caractéristiques des outils analytiques tels que les décrits J. Goody [11] , ou tels que nous apprend à les construire J. Bertin [2]. Reste, effectivement, à deviner de quelle façon une telle structuration peut influencer, par exemple, les mécanismes de la lecture et ses traditions. On a vu Manu, par exemple, parcourir des colonnes de références web en partant de la droite ou d'autres, comme Manuela, chercher à se focaliser seulement sur des espaces délimités au milieu de la page, « comme des cases ». Cette structure géographique et les stratégies d'exploration du « document » web semblent ainsi représenter l'un des traits essentiels de ce qu'après S. Broadbent on a appelé *l'architecture documentaire* numérique [4].

### 2.3 Les temporalités

Un dernier aspect du document web concerne la question de la temporalité, ou plutôt des diverses formes de temporalité associées à ses formats techniques. La question du « temps du document », et éventuellement de la façon de l'archiver, suppose en effet que l'on soit capable de traiter simultanément différents types de temporalité associés aux différents types d'objets numériques. Les animations, par exemple, sont douées d'une sorte de temporalité interne et autonome, les zones actives comme les zones de saisie ou de restitution des résultats seraient elles rythmées par le cours d'actions successives de l'utilisateur, si attentif on l'a vu aux effets mécaniques de ses manipulations. Les usagers ont aussi largement intégré le principe du « taux de renouvellement » de la page web, comme Uryèle avec ses pages du *Monde* ou de l'Académie des Pays-de-Loire, sans oublier tous ces « liens morts » qui ont pu régulièrement émailler la navigation. Associés au principe d'une mosaïque relativement flexible plastiquement, on comprend que ces modes divers de temporalité aient pu faire penser au document numérique comme une forme sémiologique « instable » [1]. Or cette « instabilité » par rapport aux cadres naturalisés du document imprimé et que souligne aussi grandement le document de R.T. Pédaque, peut aussi s'envisager comme un remarquable principe de plasticité spatiale et temporelle. On peut ainsi tirer partie du principe de « modularité » du document numérique en générant des pages « à la demande » ou sur requête : c'est le succès actuel des pages dynamiques, ce système temporaire de relations associé à un identifiant temporel entre autre. Le document, effectivement, peut devenir personnalisable et offrir des cadres variables. Mais s'agit-il encore du *même* document ? L'analyse d'utilisateurs en



contexte pose donc déjà localement la question des modes d'archivage du document numérique, de la constitution d'une mémoire documentaire sur le web et plus généralement celle d'un format temporel qui semble échapper aux catégories conceptuelles qui nous ont permis jusqu'à présent de penser l'imprimé. Que faire, par exemple, du « temps » d'un document web qui correspond seulement à celui de sa consultation locale et ne peut être reproduit, ni ailleurs, ni plus tard?

### 3 LE WEB, UN MILIEU E-COLOGIQUE

Toutes ces difficultés liées à la maîtrise des formats techniques nous a conduit, évidemment, à nous interroger sur la nature même de qu'est le web. Le terme même « d'espace » paraît peu justifié dans la mesure où il supposerait conquis un territoire qui ne se laisse pas réduire, ou une forme de système graphique ou conceptuel qui permettrait de penser le web comme un tout organisé et dont on perçoit les frontières. On pourrait alors parler d'un type inédit de milieu, un milieu *e-cologique* pour parler comme B.-A. Huberman<sup>2</sup> que les acteurs s'approprient petit à petit sans stratégie générale ou commune [12]. A vrai dire, les observés y ont plutôt évolué à tâtons, en aveugles presque, évoluant de proche en proche sans jamais voir s'incarner la promesse des *information spaces* virtuels.

#### 3.1 L'idée d'une architecture inédite

« Écran », « page », « espace d'activité », « site », « signes », « réseau », « couches »... « document » : les termes que nous avons mobilisés au cours de ces deux années manquent parfois de précision et surtout d'articulation entre eux. Ce flottement conceptuel va pourtant de soi si l'on admet le caractère radicalement inédit que forment ensemble support numérique et réseau. A vrai dire, si l'on y regarde de plus près, la distinction même entre les deux peut paraître discutable tant ils renvoient

---

<sup>2</sup>*The Laws of the Web, patterns in the ecology of information*, The MIT Press, Cambridge, Massachusetts, London, England, 2003. L'auteur est l'un des pionniers de l'étude du web à l'aide la théorie des graphes. Il considère le web comme un « gigantic informational system » qui admet de « strong regularities » comme les phénomènes de concentration de liens et de documents qui prendront le nom « d'agrégats » mais aussi des phénomènes incontrôlables et peu prévisibles comme les fameuses « Internet storms » quand on peut observer, à un endroit donné des graphes du web, la multiplication rapide des pages publiées comme une « explosion » documentaire.

de concert au même principe technique : une série de couches de calcul reliées entre elles par des opérations d'adressage. Entre « réseau » et « document numérique » il n'y a pas de différence de nature (comme pourrait nous le faire croire l'idée de la « diffusion » d'un document sur le web comme dans un milieu extérieur) mais seulement d'échelle, et encore. Le même type de « langages » informatiques les relie et assure au niveau régional d'un « document » comme au niveau global du « réseau » les relations réciproques d'une série de briques élémentaires de différentes natures. D'où la diversité à la fois sémiotique et technique des pages web : ces compositions peuvent intégrer des « pavés » de texte mais aussi des images (aux formats *gif* ou *jpeg*), des animations (type *flash*), des zones de saisie de requête et de restitution des résultats (moteurs internes, requêtes sur un serveur...), des « objets » interactifs (comme les menus déroulants). D'où aussi l'intégration de différentes données qui ne sont pas liées directement à « surface » lisible de la page : des « mots-clés » destinés aux robots des moteurs de recherche pour son indexation ou des scripts logiciels (*JavaScript* par exemple) « chargés » par le dispositif mais activés seulement à la suite de certaines actions. C'est à ce niveau que se loge le principe du lien hypertexte, pierre angulaire de toute cette architecture de la mise en relation. On pourrait encore décrire ce principe généralisé du « feuilletage », par exemple si l'on prend en compte que certaines applications logicielles peuvent se trouver directement intégrées à la page web (c'est le cas de nombreux *plug-ins* ou d'applications comme *Quick-Time*) ou de façon périphérique (on peut ainsi régler l'ouverture d'un document associé par un lien hypertexte à notre page de départ sous un logiciel comme *Acrobat*).

Un environnement dynamique et ouvert comme le web interroge donc nos cadres descriptifs traditionnels : le niveau du signe (qui mobilise une théorie de l'écriture), le texte (qui ouvre sur une théorie du document), le corpus (théorie de l'indexation et des mémoires de collections). Ces échelles pratiques pour décrire l'architecture documentaire de l'imprimé conviennent difficilement au web. En raison de sa complexité générée par l'absence d'un principe organisateur. En raison surtout de sa diversité, technique certes, mais aussi sociale et culturelle. A cet égard, le web est précisément un « milieu » où les acteurs doivent s'adapter aux contraintes multiples de l'environnement, et il est d'ailleurs souvent le territoire expérimental idéal pour toute une série d'innovations socio-techniques comme les *web logs* actuels. Ainsi s'explique la contrainte exercée sur les internautes par un milieu dans lequel coexistent une foule de « modèles » ou de « formats » documentaires. Et ils sont obligés de les maîtriser tous, sous peine d'hypothéquer la possibilité même de leur navigation. Cela s'avère d'autant plus évident si l'on s'attache à chacune des hiérarchies

construites par nos observés. Pour les uns sont des documents les articles en ligne de la presse nationale (*Le Monde*, *Libération*) ou locale (*Ouest-France*). C'est le cas pour Jeff ou Uryèle qui n'ont sauvegardé sur leurs ordinateurs que ce type de document. Pour d'autres ce sont les médias en général (sites de *ARTE* ou *M6*) ou les « sites officiels » des institutions (Ministères de la Culture, de l'Éducation). Pour les derniers ce sont les ressources web elles-mêmes, comme certains forums, listes de discussion ou sites de ressources en ligne. Et ces modèles socio-techniques de document sont soit hérités (le format *pdf* par exemple), soit inédits (hypertexte), soit en cours de construction (*wiki*, *weblogs*). C'est aussi cela pouvoir naviguer.

### 3.2 Document, médiation, mémoire

Cette diversité et cette hiérarchie des formats socio-techniques montre aussi à quel point le document est issu d'une série de médiations dont l'usager lui-même devient un rouage essentiel. Cela paraît particulièrement évident si l'on s'intéresse aux différents types de médiateurs qui lui permettent de s'orienter. En effet notre étude empirique des méthodes d'accès au Web et d'orientation dans cet univers a mis en évidence deux types de médiateurs. Tout d'abord des *médiateurs institués* dont certains n'existent que par le Web (les *Newletters* pour Julien) et d'autres ne sont que des duplications de leur rôle dans un monde social préalablement structuré (le site *Nouvelles Frontières* pour Karine). Des *médiateurs personnels* ensuite qui mettent en place une forme de coopération et de production d'orientation très locale, très adaptée à un cercle social restreint et parfois éphémère. C'est tout le rôle que jouent les proches, les amis, les relations de travail qui sont autant de relais sûrs. Ces médiateurs ne sont pas en concurrence mais peuvent tout à fait être combinés au gré des parcours d'un internaute.

La plupart du temps, c'est à partir de ces espaces du web balisés socialement que les usagers construisent différents types de *chaîne de traitement* documentaire. De l'écran à l'imprimante ou dans les dossiers sur les bureaux virtuels et réels, les univers domestiques sont en effet tissés d'un ensemble de dispositifs qui concourent à la production locale de véritables documents. À l'image d'un *centre de calcul* où s'hybrident souvent imprimé et numérique, on observe tout au long de la chaîne des opérations de « sélection, extraction, réduction » [14]. Ce qui se construit ainsi c'est un système de production de la mémoire documentaire personnelle où fonctionne un processus de filtrage de l'information, des

lieux de transit où elle se décante mais aussi où elle s'oublie, comme tous ces dossiers numériques ou imprimés « à lire » qui ne sont que rarement exploités dans leur intégralité. Elimination et enregistrement ne sont pas ici à opposer ; les deux types d'opérations relèvent ensemble d'une « écriture de la mémoire » que B. Stiegler associe à l'acte de lecture [20]. Plus qu'en termes de perte ou de gain, il s'agirait pour l'auteur de repérer le principe même de la mobilité de l'information et des supports qui préside à la construction des mémoires, cette « mobilité combinatoire des composants, la mobilité des messages sur les supports, la mobilité des supports eux-mêmes, et enfin leur reproductibilité » [20]. Cette activité de production montre que le document n'est pas seulement un ensemble de traces déchiffrées mais aussi de traces produites.

### 3.3 Un milieu pré- ou para-documentaire

La rencontre avec ce que chacun conçoit comme « document » ne représente pourtant qu'un îlot de cohérence dans un environnement expérimental sans repères majeurs ou communs. Le web nous est largement apparu au cours de l'enquête comme un espace u-topique, au double sens que lui accorde Voltaire à la fin de *Candide*. C'est un espace *idéal*, celui des « grands espaces d'information » que l'on promet chez les fournisseurs d'accès, mais qui reste seulement à l'état de promesse. En revanche c'est surtout un espace qui *n'existe pas* pour les usagers, réellement absent tant que chacun ne se le sera pas approprié pour pouvoir le partager. En somme, ce serait un milieu *pré* ou *para*-documentaire, la convergence de projets expérimentaux que l'on a du mal à s'approprier techniquement. D'où la grande diversité des formats documentaires, à laquelle certains usagers adhèrent beaucoup (« Il y a tant de choses écrites là-dedans, c'est la noyade assurée » dira Pascal), d'autres très peu, comme Laurent pour qui « il n'y a aucune référence à des choses concrètes » sur un web fait de « blabla ». Ce croisement est intéressant pour mesurer l'écart considérable avec les situations de lecture des supports analogiques tels que le livre. La maîtrise technique essentielle, celle de la lecture, est certes une condition de départ mais au-delà, la manipulation pose rarement de problèmes, à l'exception d'ouvrages spécialisés. En ce qui concerne la dimension interprétative, les formes canoniques du texte imprimé supposent une linéarité qui constitue un pré-arrangement d'un sens possible, porteur d'une intention que le lecteur finit de toutes façons par supposer et qui le contraint largement (et l'aide) dans sa stratégie de lecture. Les schémas interprétatifs ont plus de

chances de fonctionner et de s'avérer productifs, aussi singuliers soient-ils. L'objet imprimé offre donc des prises aisées dans sa manipulation (et facilitée par le format de poche ou encore par des mises en page structurées) mais le texte reste une confrontation à une altérité constituée comme telle : c'est contre elle ou à partir d'elle que le travail spécifique du lecteur peut se faire. Or c'est aussi cette expérience qu'ont parfois longuement cherché nos usagers.

## 4 DE LA TOILE AU « CHEWING-GUM »

Le web reste donc largement indéterminé du point de vue d'une offre éditoriale qui permettrait cette nécessaire « prise en charge » de l'internaute qui est parfois réclamée explicitement par les usagers. Car le web est par nature et fondation historique sans référent central, sans source ni origine, et partant sans processus de validation des pages, des sites, laissant place à la multiplication *ad hoc* des formats socio-techniques de document. C'est cette absence d'autorité qui laisse la place à la domination d'un système technique opaque, aux réactions intempêtes et aléatoires. C'est alors le système général en tant que tel qui doit d'abord être maîtrisé. D'où la personnalisation récurrente et spontanée dont il est l'objet, lorsqu'« il » ne répond pas, lorsqu'« il » ne « veut » plus afficher le texte attendu ou quand « ça fonctionne pas ». Du point de vue de l'internaute, il n'est plus possible de s'affronter à un auteur, à un texte, supposé organisé par une intention : toute intention a disparu comme principe de communication, pour être remplacé au mieux par la diversité des modèles documentaires, au pire par la simple opportunité d'un contact. C'est probablement là que se loge cette posture d'im-médiateté revendiquée qui correspond aussi à la tendance forte sur le net à la désintermédiation, à la recherche du contact direct, voire à l'autoproduction. De ce fait, il n'est pas possible non plus de développer des stratégies, ni même des tactiques, puisqu'il n'y a pas de ligne de conflit ni de principe repérable de domination ou d'asymétrie. Le contraire, pour nos usagers, du système de l'imprimé comme ensemble organisé de médiations culturelles et de ses modèles. Le contraire, aussi, du Minitel comme offre structurée où le système guide pas à pas l'utilisateur. L'imprimé a fini par produire des attentes et des règles de production qui se renforçaient mutuellement, autrement dit l'émergence d'un modèle linéaire [17]. L'offre linéaire de l'imprimé comportait à la fois une contrainte technique et une contrainte informationnelle. Dans le cas du

Web, cette offre linéaire a aussi sa place, avec ses conventions éditoriales et ses modèles documentaires (comme les sites de presse) mais elle y est marginale pour nos usagers qui n'arrivent souvent pas à se repérer comme dans un corpus préorganisé. L'action de l'utilisateur devient ici non pas possible mais *indispensable* : c'est par son activité de navigation qu'émerge un univers de connaissances qui n'avaient sans doute jamais été reliées de cette façon avant lui. Ainsi, nombre d'usagers se sentent « piégés » et ils ont toujours la hantise de « se perdre ». Dider avoue souvent « tourner en rond » et qu'il « ne s'en sort pas », tout comme Pascal pour qui « il n'y a personne pour nous sauver » car « on est seul sur ce fichu rafiote ». Nous pouvons considérer alors que le type d'emprise qu'exerce le Web et qui génère les asymétries observées et les méfiances de certains internautes, relève de la toile d'araignée, ce qui correspond bien au Web, car très éloignée de l'emprise que peut exercer un texte ou un corpus. Cependant, ce serait encore imaginer qu'existe quelque part un centre, celui où se trouve l'araignée qui tisse sa toile, et qui possède bien une stratégie pour prendre ses victimes. Or, pour le Web, il n'existe pas de centre du tout, et aucun moteur, annuaire ou portail, ne peut prétendre jouer ce rôle, même si certains le tentent ou le proclament. Nous avons plutôt affaire à un *chewing-gum*, qui n'a ni centre ni intention mais qui piège celui qui se laisse absorber, sans qu'il soit question d'y trouver des repères ou des prises pour s'en sortir.

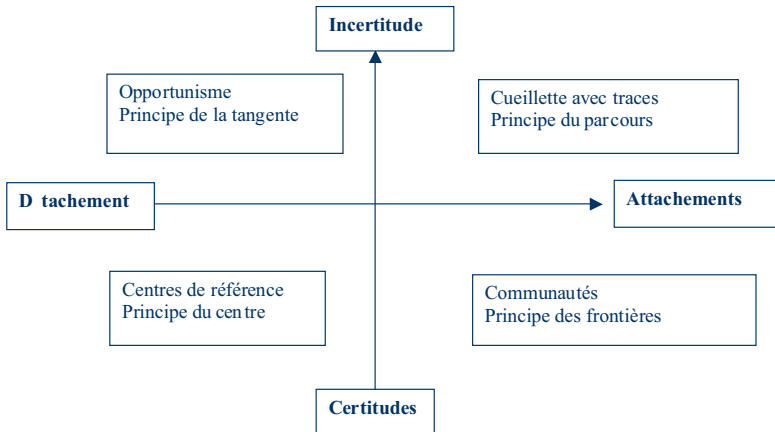
## 4.1 Topographies et styles cognitifs

Il serait donc vain de vouloir deviner sur le web l'émergence d'un seul modèle éditorial qui assurerait le cadre d'un échange (conventionnel et normé) entre un type d'offre et un type de demande. On peut même difficilement entreprendre de passer en revue pour l'heure l'ensemble des modèles documentaires présents sur le réseau étant donné leur diversité mais aussi leur manque de stabilisation dans un univers dédié pour une grande part aux expérimentations de toutes sortes. On peut en revanche s'intéresser aux façons de sortir, ou de gérer, l'incertitude native du web. Ce serait l'occasion, par exemple, de comprendre en quoi un projet comme celui du « web sémantique » peut contribuer à requalifier l'ensemble du web : ses offres, ses outils et l'ensemble des médiations qui concourraient à lui « ajouter du sens ». On peut aussi chercher, d'un autre côté, à comprendre comment les usagers eux-mêmes développent des stratégies spécifiques pour gérer cette incertitude. Dans cette perspective, on peut déjà les « classer » suivant leurs compétences dans le maniement

du système, ou même leurs « niveaux » respectifs de maîtrise. C'est ce que fait, par exemple, S. Broadbent dans son enquête, en s'intéressant plus particulièrement à ces « usagers moyens » qui feront demain selon elle le gros des utilisateurs [4]. On peut aussi se focaliser plus spécifiquement à décrire les types de parcours qu'ils effectuent sur le web comme le font Cockburn et McKenzie [5] dans une enquête qui révèle combien les pratiques du web sont routinières et épousent toujours les mêmes patterns de redondance (en moyenne, par exemple, 3 pages ouvertes sur 4 ont déjà été visitées). De là, d'autres études se développent aujourd'hui sur les stratégies spatiales de déplacement dans l'univers web, en tentant de comprendre notamment les compétences cognitives mobilisées par les usagers [8].

Nous avons plutôt choisi de chercher à identifier les stratégies d'orientation qui ont guidé les usagers en proposant quatre « modèles » ou « styles cognitifs » auxquels sont associées quatre types de topographie, en s'inspirant dans l'esprit des concepts de K. Lynch sur *l'imagibilité* de la ville [16]. Ces quatre styles cognitifs relèvent du principe général d'une *navigation sociale* que l'on aurait tort de séparer d'une simple navigation hypertexte plus « technique » [6]. Car se déplacer sur le web, c'est aussi naviguer entre différents types de transaction et de conventions et l'on comprend qu'un « expert » du web n'est pas seulement un expert en informatique mais aussi un spécialiste des postures sociales et des modes de réception.

A l'incertitude du web peut répondre une stratégie générale de construction de centres de référence partagés. Elle peut se décliner de deux façons qui peuvent se trouver d'ailleurs associées dans les pratiques. En premier lieu, un principe de construction communautaire du web permet de combiner une topographie en îlots, bien identifiés, dont on peut faire la cartographie, dont les étiquettes sont normalisées, par une autorité ou par ajustement dans le temps. Cette procédure renvoie à un style cognitif d'appropriation par territorialisation, par définition de frontières, par installation dans un espace familier. L'organisation systématique des signets (ou des favoris) peut rendre compte d'un tel travail de construction et d'orientation fondé sur des communautés de signification (supposées) partagées. En second lieu, les centres de référence partagés peuvent aussi être produits par des médiateurs réputés. Il s'agit alors d'une topographie aux repères fort visibles devenant des points de passage obligés pour s'orienter. On peut ainsi s'adosser, par exemple, aux rubriques « conseils » de son mensuel préféré pour aller voir « les bons sites ». Cette construction de la réputation peut donc être appuyée par des références acquises à l'extérieur, mais il n'est pas exclu que certains acteurs connus du web



accèdent aussi au statut de référent, comme *Yahoo/Google*. Le style cognitif correspondant se construit donc par appui sur des autorités reconnues et sur une demande et une acceptation des différentes formes de prise en charge. Dans les deux cas, il s'agit de retrouver les modes canoniques et doxiques d'organisation des savoirs, avec, par exemple, les équipements déjà présents dans les bibliothèques.

Un second type de stratégie consiste à intégrer l'incertitude du web dans ses règles d'action et comme cadre interprétatif. La perspective adoptée est en quelque sorte beaucoup plus relativiste et centrée sur soi comme centre de référence. Dans un premier cas, on peut ainsi opérer par capitalisation de traces successives en se servant de ses propres parcours de navigation comme d'une forme mémoire. C'est tout le principe de ce l'on appelé par ailleurs «indexation subjective» et que l'on peut qualifier de «cueillette avec traces». Dans cette optique, tout est bon à enregistrer et ce n'est souvent que rétroactivement que s'organise la cohérence des traces accumulées. On comprend dans ce cas le succès d'outils comme *Copernic* ou *Nestor* qui conservent, l'un sous forme de tableau, l'autre sous forme de graphe, les traces d'une recherche et des sessions de navigation sur le web. Dans un second cas, cette démarche relativiste peut déboucher sur une stratégie du détachement et de l'opportunisme. Ce serait le cas des (vrais) surfers qui acceptent l'absence de centres et de repères d'un réseau peu structuré sémantiquement, avec sa faible traçabilité et l'indétermination des liens qui l'organise. Ne recueillant que peu de traces, ils veillent les occasions et sont remarquablement prêts à jouer de nombreux rôles de façon garder fluide leur navigation. Paradoxalement, c'est ce type de stratégie qui correspond peut être le mieux à ce qui semble être la nature du web et qui produit donc les plus



fameux experts.

Nous pouvons représenter ces quatre stratégies sous la forme d'un graphe organisé selon la tension entre attachements [15] et détachement qui est l'enjeu du conflit entre tradition et modernisme d'une part et entre certitudes et incertitude qui devient l'enjeu du nouveau basculement depuis le modernisme vers l'hypermodernisme ou le relativisme d'autre part. Tous ces positionnements permettent de relier ces pratiques personnelles aux débats actuels sur le Web sémantique par exemple, sur le poids des majors de l'édition, ou encore sur la navigation sociale comme modèle possible d'organisation émergente.

On notera, pour conclure, que ces stratégies pour réduire la diversité des modèles d'échange documentaire (quand ils sont présents) et pour lutter contre l'indétermination générale du réseau ont un « coût » cognitif non négligeable. Certains usagers, en effet, n'hésitent pas à interrompre une recherche, un parcours ou une lecture quand ils se sentent « perdus » techniquement par un système qui semble les dépasser, ou socialement par le foisonnement des références à maîtriser pour s'y repérer. C'est dire à quel point la navigation sur le web réclame aussi une forme de réflexivité sur les modèles interprétatifs mobilisés et leur relativité, ce que la plupart des usagers observés au cours de ces deux années ne sont pas encore enclins à faire.

## Références

- [1] Synthèse établie par Roger T. Pédaque dans le cadre du réseau thématique pluridisciplinaire 33 du département STIC du CNRS, juillet 2003.
- [2] J. Bertin, *la sémiologie graphique*, Gauthiers-Villars - Mouton, Paris - La Haye, 1967.
- [3] D. Boullier. *L'urbanité numérique. Essai sur la troisième ville en 2100*, Paris : L'Harmattan, 1999.
- [4] S. Broadbent, S. Cara, «A Narrative Approach to User Requirements For Web Design», *Interactions*, Dec. 2000, *ACM Press*.
- [5] A. Cockburn, B. McKenzie, « What Do Web Users Do ? An Empirical Analysis of Web Use », *Int. Journal of Human-Computer Studies* (2000).
- [6] Dieberger, A., Dourish P., Höök, K., Resnick, P. and Wexelblat, A., 2000. Social Navigation. Techniques for Building Mode Usable Systems, *Interactions*, vol. VII, n°6, 36-45.
- [7] F. Ghitalla, D. Boullier, P. Gkouskou, L. Le Douarin, A. Neau. *L'outre-lecture. Manipuler, (s') approprier, interpréter le Web*, Paris : Bibliothèque

- Publique d'Information/Centre Pompidou, 2003, 267p.
- [8] F. Ghitalla, C. Lenay, « Les Territoires de l'information. Navigation et construction des espaces de compréhension sur le web » in *La Navigation*, Les Cahiers du Numérique, vol.3, n°3-2002, p.51-63.
- [9] F. Ghitalla. « L'espace du document numérique », *Communication et Langages*, n°126, A. Colin, 2000 ; « NTIC et nouvelles formes d'écriture », *Communication et Langages*, n°119, Retz, 1999.
- [10] C. Ginzburg, *Mythes, traces, emblèmes. Morphologie et histoire*, Flammarion, 1986.
- [11] J. Goody. *La raison graphique*. Paris : Eds de Minuit, 1979.
- [12] B.-A. Huberman. *The Laws of the Web, patterns in the ecology of information*, The MIT Press Cambridge, Massachussets, London, England, 2003.
- [13] Y. Jeanneret et E. Souchier, « Pour une poétique de l'écrit d'écran », *Xoana*, n°6-7, le multimédia en marche, J.-M. Place, 1999.
- [14] B. Latour, « Ces réseaux que la raison ignore : laboratoires, bibliothèques, collections », in *Le Pouvoir des Bibliothèques. Le pouvoir des livres en Occident*, sous la dir. M. Baratin et Ch. Jacob, Albin Michel : Paris, 1996, p.23-46.
- [15] B. Latour. *Politiques de la nature. Comment faire entrer les sciences en démocratie*, Paris : La Découverte, 1999.
- [16] K. Lynch. *L'image de la cité*, M.I.T. Press, Paris : Dunod, 1976.
- [17] Manovich, L. *The language of new media*, Cambridge, Ma, MIT Press, 2000.
- [18] D. Olson, .- *L'univers de l'écrit*, Paris : Editions Retz, 1998 ( *Worlds of paper*, Cambridge University Press, 1994)
- [19] E. Souchier, « L'écrit d'écran. Pratiques d'écriture et informatique », *Communication et Langages*, n°107, A. Colin, 1996.
- [20] B. Stiegler, « Machines à lire », in *La Bibliothèque, miroir de l'âme, mémoire du monde*, Autrement, série « mutations » n°121, Paris, avril 1991.

