

Un objet hybride : le document numérique

Stéphane Caro

Dans quel cas a-t-on à faire à un document numérique ? Certaines pages *web* sont constituées de façon dynamique et n'ont d'existence que le temps d'une consultation. La page 'actualités' de *Google* est mise à jour toutes les quinze minutes par un moteur de recherche qui balaye 500 sites d'informations journalistiques¹. Lors d'une prochaine visite de cette page, son aspect sera différent. Le document numérique transforme les logiques de définition, de légitimité et de stabilité des documents sur support papier².

Dans le cadre de cet article, nous définirons le document numérique comme un document dont le mode de consultation principal est la lecture à l'écran selon la formule de D. Cotte : *Il convient d'isoler, ou de réserver le terme de document numérique uniquement aux documents ou à la phase d'édition des documents qui ne retrouvent pas une forme « analogique » mais subsistent sous une forme consultable au moyen d'un appareillage électronique*³. Par extension, certaines formes d'affichage sur écran peuvent s'apparenter en ce cas à des documents numériques même si elles n'ont pas un objectif de trace ou de stabilité (écrans de logiciels, de tableaux de contrôle de processus). La frontière est en effet désormais très ténue entre document et logiciel. Certains sites *web* par exemple proposent une partie que l'on peut qualifier de *document* tandis qu'une autre partie s'apparente davantage à un *logiciel*. La société Bull a ainsi développé des logiciels sous forme de site Internet qui permettent à distance le contrôle et l'administration de serveurs informatique. S'agit-il d'un document ou d'un logiciel ? Le monde numérique pose donc la question de l'émergence d'un nouveau type de document dont le contenu peut être instable et varier notablement d'une consultation à l'autre du fait de l'intervention d'applications informatiques dans la détermination d'une partie du document (bases de données, programmation, animations). Une définition ancienne du « document » de Robert Escarpit, au centre de sa théorie générale de l'information communication, doit être repensée du fait de l'instabilité des documents numériques et de leur immatérialité : *Un objet informationnel visible ou touchable et doué d'une double indépendance par rapport au temps* :

- synchronie : *indépendance interne au message qui n'est plus une séquence linéaire d'événements, mais une juxtaposition multidimensionnelle des traces ;*
- stabilité : *indépendance globale de l'objet informationnel qui n'est plus un événement inscrit dans l'écoulement du temps, mais un support matériel de la trace qui peut être conservé, transporté, reproduit.* »⁴.

L'instabilité des documents numériques est plurielle : topologique pour les documents en réseau (localisation de la ressource informatique, de la page dans le site), discursive (des parcours de

¹ J.-M. Salaün, « Chronique inachevée d'une réflexion collective sur le document », dans *Communication & langages*, n° 140, Juin 2004, p. 9-17.

² R.-T. Pédaque, *Document : forme /archives, signe et relation, les re-formulations du numérique. Document de travail*, STIC-CNRS. En ligne http://www.archivesic.ccsd.cnrs.fr/documents/0/00/00/04/13/index_fr.html. 20 p.

³ D. Cotte, « Le concept de 'document numérique' », dans *Communication & langages*, n° 140, juin 2004, p. 31-41.

⁴ R. Escarpit, *L'information et la communication : théorie générale*, 1876, Hachette, Paris, rééd. 1991, p. 123.

lecture possible), du contenu qui est mis à jour et enrichi continuellement⁵. Les définitions successives du document vont de ce fait vers une intégration de multiples dimensions qui ne se limitent plus au contenu et au support mais intègrent l'auteur, le contexte d'élaboration et de réception ainsi que d'autres variables⁶.

Une esquisse de définition récente du concept de document a été proposée par un collectif de chercheurs dans le cadre de réflexions sur le document numérique :

*« Un document ne serait finalement qu'une convention entre des hommes qui formerait une part de leur humanité, de leur capacité à vivre ensemble et dont les modalités anthropologiques (lisibilité perception, signe), intellectuelles (intelligibilité-assimilation, texte) et sociales (sociabilité-intégration, médium) devraient non seulement être efficaces prises chacune séparément, mais encore être cohérentes entre elles ».*⁷

Cette définition illustre la difficulté posée par l'émergence des documents numériques quant à la définition même du concept générique de document.

Quelques caractéristiques du document

Un document peut se décrire à l'aide de nombreux facteurs (support, accès, discours etc). Une des caractéristiques principale des documents numériques est leur interactivité (utilisation des liens) et la présence de différents types d'informations ou plus précisément « modes de communication ». Ces modes de communication sont tantôt statiques : textes, photographies, illustrations, dispositifs de présentation synoptiques (diagrammes, tableaux, organigrammes etc.) et tantôt dynamique, intégrant une dimension temporelle, le son (la voix, la musique...), l'image animée (film, animation etc.). L'interactivité et la présence de modes de communication de divers type, avec ou sans dimension temporelle, regroupés dans un même document et stockés sur un espace unique est particulièrement caractéristique des documents numériques⁸.

Interactivité

L'hypertexte est le mode de gestion de l'information privilégié des documents numériques. Dans ce mode, l'information est représentée sous la forme de blocs de textes ou unités d'informations, appelés nœuds⁹. Ce mode de fonctionnement et l'immatérialité des documents engendrent des difficultés d'usage souvent mentionnées comme :

- l'abstraction qui rend difficile l'appréhension du document par les utilisateurs : difficultés de représentation de la structure, du contenu du document et de son volume.
- la « surcharge » cognitive liée aussi en partie à l'abstraction de nombreuses dimensions du document. Ceci entraîne parfois une charge mémorielle importante. Pour une recherche d'information par exemple, il faut mémoriser l'information recherchée ainsi que son parcours ou la structure du document, sa position courante.

⁵ B. Juanals, *La culture de l'information : du livre au numérique*, Hermès, Paris, 2003.

⁶ V. Couzinet, « Le document : leçon d'histoire, leçon de méthode » dans *Communication & langages*, n° 140, juin 2004, p. 19-29.

⁷ J.-M. Salaün, « Chronique inachevée d'une réflexion collective sur le document » dans *Communication & langages*, n° 140, juin 2004, p. 9-17.

⁸ M. Melot, « Le multimédia : un objet traditionnel » dans *Degré*, n°92-93, hiver 1997/printemps 1998.

p b2 à b12 et P. Lévy, *Cyberculture*, Odile Jacob, Paris, 1998.

⁹ M. Nanard & J. Nanard, *Documents hypermédiés : Définitions et concepts associés. Techniques de l'ingénieur, traité informatique*. Techniques de l'ingénieur, Paris, 1999, 22 p.

La concaténation des documents numériques en réseau aboutit à une forme d'espace informationnel sans limites « dont le centre est partout et la circonférence nulle part » sorte d'ordinateur « hypertextuel, dispersé, vivant, pullulant, inachevé », appelé *Cyberspace*¹⁰.

Le terme multimédia à plusieurs sens. Support éditorial (journaux, TV, radio...), support physique de stockage (DVD, cédérom...) et mode de communication (son, image, texte). Ce dernier sens semble être adopté pour désigner l'intégration de modes de communication dans des techniques de reproduction, transmission et de stockage. La multimodalité serait la possibilité d'interagir avec plusieurs sens (toucher, odorat, goût). Il y a multimodalité lorsque plusieurs canaux perceptifs sont utilisés en synergie, parallèlement ou séquentiellement, pour convoier la totalité d'un message¹¹. Typiquement lorsqu'un humain indique oralement « mettez cet objet ici » en désignant avec le doigt l'endroit où déposer l'objet, il s'agit de communication multimodale. De même il est possible de sélectionner un objet virtuel avec un pointeur comme la souris d'un ordinateur par exemple et de lui appliquer une transformation à l'aide d'une commande vocale. Chaque canal convoie une partie de l'information. Du point de vue des interactions, certains auteurs évoquent le bi-média car « multi » se résume en fait souvent à la sollicitation de deux canaux et à deux modes seulement de communication : sonore et visuel. Si l'on s'attache à caractériser le support de l'information, on peut aussi rencontrer le terme d'unimédia, l'information étant enregistrée sur les supports informatiques en 0 et 1. Le terme multimédia a donc un sens différent selon le point de vue retenu.

Terminologie du découpage de l'information

La question du découpage de l'information en vue de la mise en écran est centrale et ancienne puisqu'elle est propre à la présentation d'informations dans un espace limité (Comme les tablettes d'argile sumériennes). Cette question du découpage de l'information en vue de la structuration globale du document et de la mise en écran amène de nombreuses appellations des éléments ainsi obtenus. Balpe distingue par exemple trois niveaux de segmentation de l'information dans un document numérique¹². Ce qui relève du thème général du document ou le « domaine » : (les vins, la taille des rosiers, la documentation technique d'un avion etc). Les « unités d'informations » qui correspondent aux nœuds d'information dans l'environnement hypertexte, soit sensiblement l'espace d'un écran. Enfin les concepts qui sont des sous parties des unités d'informations correspondant chacun à l'expression d'une idée, (de l'ordre du paragraphe).

D'autres termes ayant des significations similaires sont rencontrés dans la littérature :

- nœud ou unité d'information : généralement équivalent à un écran ou une page (*Sites Web*). Cet espace doit être auto-compréhensible car il peut être consulté de façon indépendante¹³.
- concept, unité textuelle : généralement équivalent à un paragraphe ou un acte de langage spécifique. On place plusieurs unités textuelles ou concepts sur un écran ou une page (*sites web*).
- domaine : ensemble des connaissances sur un sujet dont fait état le document numérique. Exemple : les recettes de cuisine bourguignonnes.

¹⁰ P. Lévy, *op. cit.* p. 45.

¹¹ M. Nanard & J. Nanard, *op. cit.* p. 3.

¹² J.-P. Balpe, *Hyperdocuments hypertextes hypermédiés*, Eyrolles, Paris, 1990, 200 p.

¹³ M. Nanard & J. Nanard, *op. cit.*

– para-texte¹⁴, unité textuelle de second plan¹⁵, désignent des unités textuelles périphériques accessoires à l'objet central du document.

D'où vient cet objet ?

Il est fréquent dans la littérature de faire remonter l'histoire des documents numériques à l'histoire des mécanismes liés au fonctionnement des documents numériques. Certains auteurs estiment en effet que les concepts élémentaires des documents numériques sont très anciens. Ainsi l'espace contraint de l'écran est parfois comparé aux tablettes d'argile sumériennes, le défilement du texte à l'écran au défilement des manuscrits enroulés. L'espace multiple de la page aux bibles commentées et illustrées du moyen âge. Ce qui est foncièrement nouveau c'est l'association de ces concepts et de nouveaux dispositifs au sein d'un document unique. Le document numérique s'inscrit dans une volonté d'intégration de modes de fonctionnements anciens et des modes de communication à leur support¹⁶. Il semble que le document numérique commence à émerger avec le traitement de texte et la convergence entre matériel informatique et matériel de secrétariat (la bureautique). Avant que ce traitement ne soit informatique ce traitement a été mécanographique ou microphotographique dans des projets ambitieux qui n'ont pas toujours vu le jour Otlet (1935)¹⁷, Bush (1945). Le projet *Memex* de Bush prévoyait d'automatiser des relations entre textes de façon analogue à des associations d'idées d'un être humain. L'hypertexte va révolutionner l'approche du document en réseau en particulier à travers les sites Internet¹⁸. Par la suite dans les années 1960-70 le traitement des textes s'est informatisé à des fins d'étude des textes (analyse de discours) de traduction de recherche d'information, de reconnaissance de l'écriture, et pour la photocomposition¹⁹. Dans les années 70-80 l'émergence de la micro-informatique conduit à l'apparition de traitements de textes destinés à simplifier les tâches de bureau. Les traitements de textes seront tantôt *wysiwyg* (*What You See Is What You Get*) ce qui s'affiche à l'écran représente ce que l'on obtiendra à l'impression ou *batch*, la mise en page est alors produite seulement lors de l'impression²⁰. Les interfaces graphiques ou WIMP (*Windows, Icons, Mouse, Pulldown menus*, fenêtres, icônes, souris, menus déroulants) permettront l'utilisation de traitements de textes *wysiwyg* (*What You See Is What You Get*). Les systèmes de traitement de texte apportent une première fracture entre forme et contenu. Le texte est dissocié de règles de présentation qui vont être appliquées pour la visualisation ou l'impression. Les années 80 verront l'émergence du « document structuré » et de la manipulation de documents avec de nouvelles caractéristiques prêtées aux documents comme l'interactivité, l'insertion de nouveaux modes de communication dans le document (images, formules mathématiques...) mais aussi de balises de marquages des différents types d'informations contenues dans le texte. Depuis les années 1990, le document se dote de modules logiciels qui en font des documents de plus en plus interactifs et hybrides entre applications logicielles et document numérique. Actuellement c'est le document en réseau qui se développe avec toujours un aspect dynamique notamment pour les documents reliés à une base de données, dont le contenu peut varier au fil des consultations.

¹⁴ G. Genette, *Palimpsestes : la littérature au second degré*, Le Seuil, coll. Points Essais, Paris, 1997.

¹⁵ S. Caro & A. Bisseret, *Étude expérimentale de l'usage des organisateurs para-linguistiques de mise en retrait dans les documents électroniques*, dans *Le Travail Humain*, tome 60, n° 4, 1997, p. 409-437.

¹⁶ M. Melot, *op. cit.*

¹⁷ En 1934, Paul Otlet publia à Bruxelles le *Traité de documentation* contenant entre autres des visions prophétiques sur l'utilisation documentaire des télécommunications, l'anticipation du principe d'hypertexte dans la description d'un système d'information et la prémonition d'un réseau international d'échange de données...

¹⁸ J. Nanard & M. Nanard, *op. cit.*

¹⁹ J. André, *Une vision historique du concept de « document numérique »*. Document en ligne : <http://www.irisa.fr/faqtypo/histdoc.pdf>, 15 p., consulté le 10 mars 2003.

²⁰ *Ibid*

Conclusion

Le document numérique conduit à repenser le concept générique de document. La définition de plus en plus large de « document » circonscrit difficilement les contours de plus en plus flou d'un objet aux caractéristiques mouvantes, d'un point de vue technique et rhétorique. Les outils d'interaction n'ont pas terminé leur évolution. Il n'existe pas à proprement parler de rhétorique arrêtée de l'écriture à l'écran, cette dernière est en effet inféodée aux évolutions des dispositifs techniques. La frontière entre application logicielle et document tend à s'estomper ce qui présage d'un futur incertain quant aux compatibilités des dispositifs de consultation dans les années à venir. Une des idées force qui a servi l'essor du document numérique en réseau sur Internet, la possibilité de consulter les premiers sites Internet sur tout type de plateforme avec n'importe quel navigateur est en passe de s'effondrer. À grand renfort de rustines logicielles et de solutions propriétaire, les dispositifs de consultation ad hoc pour documents numériques hybrides diminuent les possibilités d'accès aux documents numériques déjà limitées par les barrières de la fracture numérique. L'objet « document numérique » a des contours flous et il faudra encore du temps pour voir émerger une définition précise de cet objet ainsi qu'une rhétorique de l'écriture de ces documents. Ces problématiques nourriront de nombreuses pistes de recherche dans plusieurs disciplines qui étudient cet objet.