



HAL
open science

Forum DocNum - présentation des textes et commentaires

Marc Augier, Evelyne Broudoux, Dominique Cotte, Olivier Ertzscheid,
Gabriel Gallezot, Yves Jeanneret, Sylvie Leleu-Merviel, Niels Windfeld Lund,
Jean-Hugues Réty, Jean-Michel Salaün, et al.

► **To cite this version:**

Marc Augier, Evelyne Broudoux, Dominique Cotte, Olivier Ertzscheid, Gabriel Gallezot, et al.. Forum
DocNum - présentation des textes et commentaires. 2004. sic_00001099v2

HAL Id: sic_00001099

https://archivesic.ccsd.cnrs.fr/sic_00001099v2

Submitted on 11 Mar 2009

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.

Les auteurs

Texte d'introduction

à la discussion :

[Salaun J.-M.](#)

Textes à discuter :

[Cotte D.](#)

[Gallezot G.](#)

[Jeanneret Y.](#)

[Merviel S.](#)

Les commentaires

sur les textes

[Niels W.L.](#)

[Totschnig M.](#)

[Zeller J-D](#) [1] [2] [3]

[Broudoux E.](#) [1] [2]

[Augier M.](#)

[Réty J.-H.](#)

[Ertzscheid O.](#) [1] [2] [3] [4]

Pour une collaboration STIC - SIC sur le "Document"

Jean-Michel Salaun

Le département Sciences et technologies de l'information et de la communication (STIC) du CNRS a mis en place une série de réseaux thématiques pluridisciplinaires. Parmi ceux-ci, il en est un intitulé « Documents et contenu : création, indexation, navigation » (en abrégé : RTP-DOC [1]). Sa thématique recoupe ce que les anglo-saxons appellent « Information science », regroupant principalement les informaticiens qui se préoccupent du document électronique et les chercheurs de la sensibilité plus « informationnelle » des Sciences de l'information et de la communication telle qu'on les pratique dans la 71e section des universités françaises.

Pour engager une réflexion collective entre des chercheurs qui souvent ne se connaissaient pas ou mal, nous avons utilisé pleinement les potentialités de la structure en réseau et celles des échanges électroniques. L'objectif, clairement exprimé dès le départ du travail, quelques mois après la fondation du réseau en mai 2002, était la rédaction d'un texte collectif sur le "document".

Voici, brièvement résumées [2], les différentes étapes du processus de rédaction, tel qu'il s'est réalisé au jour où nous écrivons ces lignes sachant qu'il est encore en cours et que nous ne pouvons dire quand (ou même si) il s'achèvera :

- une première phase préliminaire de contributions et discussions libres, soit par la liste générale du réseau, soit dans ses instances de pilotage, s'est conclue par un séminaire qui a réuni une trentaine de chercheurs. Le résultat de cette première phase a été un balisage des approches et recherches en cours. Sur la base de tous les matériaux recueillis, j'ai synthétisé un premier texte qui a été signé d'un pseudonyme : Roger T. Pédaque (pour RTP-DOC). S'agissant d'un document de travail reprenant de nombreuses contributions de natures diverses, il était difficile de lui attribuer une responsabilité auctoriale précise. Ce choix, au départ pragmatique et amusé, s'est avéré d'une réelle efficacité. Il a en effet rendu caduque toute présomption de propriété intellectuelle sur tel ou tel argument du texte et a donné d'emblée une consistance symbolique à la dynamique en l'incarnant dans l'épopée d'un personnage fictif.
- La véritable rédaction collective s'est développée dans la deuxième phase. Le texte du susdit Roger a été mis en ligne sur un site d'archives ouvertes [3]. Chaque paragraphe était numéroté et il était indiqué que le lecteur pouvait participer à sa critique et donc à son amélioration en envoyant ses remarques à l'adresse de l'auteur, précisant le numéro du ou des paragraphes incriminés. Parallèlement, une liste de diffusion dédiée était créée pour recueillir et partager les différents apports.

Les membres du comité de pilotage du réseau étaient inscrits à cette liste ainsi que, au fur et à mesure, les différents intervenants sur le texte. Une large publicité sur l'opération a été faite à la fois à l'intérieur du réseau et dans diverses listes spécialisées ou même individuellement auprès d'experts du domaine. La discussion a été ouverte pendant deux mois. Elle a rencontré un vrai succès puisque le texte a été téléchargé plus de mille fois et qu'une trentaine de chercheurs ou professionnels du document ont proposé des corrections, amendements ou commentaires. Ceux-ci étaient intégrés au fur et à mesure dans une version électronique du texte sous forme d'annotations [4]. La conclusion de cette phase a été, évidemment, une nouvelle version, enrichie et corrigée, du texte qui fut assez simple à reconstruire compte tenu du dispositif choisi. Cette seconde version a été mise en ligne début juillet 2003 sur le site d'archives et est encore, au moment où nous écrivons ces lignes, la version de référence [5].

- La troisième, et dernière à ce jour, étape est de retrouver le processus scientifique traditionnel pour tirer un véritable profit de recherche de la dynamique lancée. A cette fin, le texte de Roger T. Pédaque est pris comme point de départ à des appels à contribution [6], des dossiers de revues, des débats électroniques ou réels, mais cette fois les articles sont signés par leurs auteurs, révisés par les pairs, et les responsabilités éditoriales sont assumées. Le point d'orgue de l'opération sera la semaine du document numérique qui se tiendra du 21 au 25 juin 2004 à La Rochelle [7]. Au moment où j'écris ces lignes, cette étape est en cours et il n'est pas possible d'en anticiper les résultats.

J'ai eu l'occasion de présenter les travaux de Roger T. Pédaque au colloque de la SFSIC de Bucarest à l'automne 2003. C'est ainsi que l'idée d'une discussion sur ce texte, mais initiée par la SFSIC est née. Au sein des Sciences de l'Information et de la Communication (SIC) bien des chercheurs et des problématiques s'attachent au document, soit frontalement, soit à l'occasion de travaux et réflexions connexes, et ceci largement au delà des dites "Sciences de l'information". Mais, soit pour des raisons de géographie institutionnelle, soit à cause d'une méfiance ancienne, soit simplement par méconnaissance, manque de temps et de disponibilité, ou enfin par une posture volontairement distanciée, peu de contributeurs issus des SIC ont jusqu'ici participé au travail collectif du RTP.

Pourtant les SIC étant ma communauté d'origine, je sais bien que ses chercheurs auraient beaucoup à dire ou à redire sur le sujet. Je sais aussi que leur contribution est essentielle pour une analyse lucide des mouvements qui aujourd'hui concernent le document. Et, maintenant que je connais de l'intérieur une partie de la communauté STIC, je suis persuadé de l'importance du dialogue qui doit se nouer, dans le respect des positionnements scientifiques des uns et des autres.

C'est pourquoi je me réjouis de l'initiative de ce forum. Je suis curieux et intéressé par ses

résultats, autant sur le fond, la discussion scientifique, que sur la forme, l'animation et l'hypermédia qui doit en découler.

Notes :

1. Pour une information sur l'ensemble des activités du réseau : <http://rtp-doc.enssib.fr>.
2. Pour une présentation documentée de la réflexion collective : <http://rtp-doc.enssib.fr/pedauque/index.html>
3. Document : forme, signe et relation, les re-formulations du numérique, Roger T. Pédaque. Article. 28 avril 2003. Working paper. http://archivesic.ccsd.cnrs.fr/sic_00000413.html
4. L'ensemble des contributions intermédiaires et le détail de la démarche est consultable sur le site du RTP à la rubrique « Roger T. Pédaque ».
5. Document : forme, signe et médium, les re-formulations du numérique, Roger T. Pédaque. Article. 08 juillet 2003. Working paper. http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/05/11/index_fr.html
6. Pour la revue I3 <<http://www.revue-i3.org>>
7. <http://sdn2004.univ-lr.fr/>

Le concept de « document numérique ».

Dominique Cotte

[DC-1]

[DC-1] La notion de « document numérique » a le mérite de re-questionner le concept même de document, à partir d'une réflexion sur le rôle de la technique dans la fabrication des signes et des supports de communication. Mais si la part technique de l'énonciation éditoriale, longtemps négligée, voire méprisée, mérite pleinement d'être prise en compte, on ne saurait considérer le document sous ce seul aspect.

[DC-2] Plus qu'une nouvelle catégorie, définitive, de documents qui s'opposerait à celle du document analogique ou « papier », le caractère numérique est ici présenté comme un état, parfois transitoire, parfois définitif, mais toujours réversible et transformable. Du même coup, il importe de ne pas tomber dans des jugements définitifs sur les qualités du document numérique, mais au contraire de prendre en compte les évolutions et métamorphoses qui se produisent lors de ces changements d'état.

[DC-3] Alors que les grands calculateurs travaillant sur des données datent de la fin des années 1940 et que les techniques du « traitement de texte » via la photocomposition investissent le secteur de l'édition et de la presse à partir du milieu des années 1970, on ne parle du « numérique » que depuis une dizaine d'années seulement, à partir de la francisation du vocable anglo-américain digital. Plusieurs points sont à remarquer ici.

[DC-4] Tout d'abord, l'objet « numérique » est un objet qui ne dit pas son nom. Il serait plus juste de parler d'objet informatique ou plus exactement électronique, car il relève d'une technologie qui repose sur une organisation binaire : le passage ou non du courant électrique. Comme ces états d'affirmation et de négation se traduisent symboliquement par le 0 et le 1 qui sont des nombres, et que par ailleurs on arrive ainsi à réduire une collection significatives de signes différents à des objets de calcul, un glissement s'est opéré progressivement qui finit par faire de la base technique le caractère principal d'objets dont les formes et la nature au plan de la communication est par ailleurs parfaitement hétérogène.

[DC-5] Il n'y a pas lieu ici de refaire l'historique d'un terme, encore que cette dimension épistémologique nous paraisse un détour nécessaire chaque fois qu'il est question d'évoquer des notions attrape-tout, comme le « multimédia ». Néanmoins, on pourra se rappeler que c'est à partir du son, et de la mise au point par Philips et Sony au début des années 1980 (après des années de recherche) du compact disc audio, que la notion de numérique commence à s'imposer au grand public.

[DC-6] Les NTIC, technologies numériques, et les réseaux comme internet

accentueront le phénomène, jusqu'à l'explosion, au début des années 2000 de l'image numérique avec la vidéo et la photo grand public. Les sites de commerce électronique, les grandes surfaces, la presse, spécialisée ou non, y vont de l'habituelle injonction, il s'agit de « passer au numérique », selon la logique qui veut que quiconque ne s'inscrit pas dans les canons de la modernité technologique est frappé de ringardise ou d'ostracisme médiatique.

[DC-7] Sous des dehors apparemment objectifs, puisqu'il est formé à partir de sa base technologique, le vocable numérique est pourtant porteur de nombreuses représentations idéologiques, puisqu'il voisine avec virtuel, immatériel, intuitif, transparent. Ce ne sont plus des objets qui changent de substrat technologique, mais un univers entier qui se présente à nous sous une forme où bien sûr, les médiations, les difficultés de la construction et de la circulation des savoirs seront forcément aplanis, puisque tout sera unifié dans un même format.

[DC-8] On reconnaît là les utopies de la communication qui marquent chaque innovation technologique [1]. Il reste que ce mouvement de transfert d'une forme de reproduction technique à une autre change malgré tout la forme, le statut et les conditions de fabrication et de réception de ces objets que sont les textes et les documents ; objets qui permettent la relation entre les hommes. Le fait que le mot « document » ait besoin d'être qualifié par l'adjectif numérique pour que soit signifié son changement de statut finit par reposer la question de la nature même de cet objet, dont on verra qu'elle est loin d'être aussi sûrement définie.

[DC-9] Parler de document numérique, c'est insister sur la facture technique, au détriment des autres caractéristiques. Au début des années 1990, une chercheuse américaine, Carolyn Marvin, livre une enquête historique fort intéressante sur le foisonnement technologique du dernier quart du 19^e siècle sous le titre « When old technologies were new » [2].

[DC-10] L'ouvrage est sous-titré : « Thinking About Electric Communication in the Late Nineteenth Century » et effectivement Marvin ramène des technologies aussi différentes que le téléphone, le phonographe, la radio, à un même substrat technique qui est l'électricité. Si l'on y réfléchit bien, il serait tout à fait légitime, en ce sens de qualifier n'importe quel document de document électrique, puisqu'il a fallu recourir à de l'énergie pour le produire. Or, cela n'a pas de sens, car ça ne caractérise en rien la nature d'un tel document.

[DC-11] Insister sur les origines techniques du document numérique ne suffit pas en soi à le distinguer. Les documents, comme tout objet produit par l'activité humaine, relèvent d'une dimension technique. Ils sont le produit de la rencontre entre un agent, un matériau et un outil de travail.

[DC-12] Le terme « numérique » dans la généralité de sa promesse technologique, ne précise rien. Aujourd'hui, tout document produit, ou presque, tend à connaître un état transitoire sous forme de document numérique. Mais cette nouvelle phase

d'industrialisation du process de fabrication n'induit pas forcément des changements dans l'état final du document, tel qu'il est rendu au lecteur, à l'auditeur, au spectateur pour être consommé. Un média comme la radio par exemple, témoigne d'une remarquable stabilité de forme pour l'auditeur, alors même que tout le processus de fabrication des émissions et surtout d'archivage est en voie de bouleversement total.

[DC-13] Le journal imprimé, comme le livre, sont loin d'être condamnés, par ce qu'on a appelé les « nouveaux médias ». En revanche, il conviendrait 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. Il en est ainsi des cédéroms et des sites web par exemple.

On avance d'ordinaire un certain nombre de traits distinctifs pour caractériser le « document numérique ».

Ouverture.

[DC-14] Le document traditionnel, à l'issue de son processus de fabrication, est livré dans une forme rigide, définitive, qui ne laisse pas d'autre place à l'intervention du lecteur qu'un ajout de l'extérieur, une empreinte supplémentaire qui ne modifie pas le cœur du document. Un livre peut être corné, annoté, souligné, surligné, mais cette intervention du lecteur n'est qu'un dialogue différé, amoindri et pour tout dire vain avec l'auteur.

A l'inverse, un fichier texte, qui serait, dans une logique traditionnelle, une version préliminaire du document, peut transiter par les réseaux comme le document lui-même. Il est, sous cette forme, susceptible de subir toutes les altérations, voulues ou non.

[DC-15] Le numérique ne produit pas ces altérations, car on ne l'a pas attendu pour fabriquer des faux, des écrits apocryphes, des contre-façons, des témoignages truqués, mais il en facilite considérablement la manipulation. En recopiant des passages entiers de livres dans ses copies, le cancre analogique en profitait sans doute pour mémoriser tant bien que mal quelques notions, tandis que le cancre high tech peut se contenter de connaître les manipulations de l'architexte sous la forme du couper-coller.

Automaticité.

[DC-16] Objet de calcul, le document numérique est paramétrable. Son découpage, classique, en zones logiques et signifiantes comme le titre, le texte, les paragraphes est également susceptible de fournir matière à un re-découpage permanent pour les

besoins de l'édition multimédia par exemple, ou pour la création de méta-documents ou de méta-textes comme dans le cas des listes-résultats produites par les moteurs de recherche.

Plasticité et recomposition.

[DC-17] Ces caractéristiques découlent directement de ce qui précède. On a l'idée prégnante que le document s'est émancipé de son support. Qu'il est un ensemble de pièces, re-combinable à l'infini, et que en tant que tel, il échappe à sa définition classique. Mais il y a probablement là en partie une illusion ; même si cette possibilité intrinsèque introduit une dimension radicalement nouvelle, il convient d'être prudent dans l'analyse de ses effets.

[DC-18] Premièrement, le passage d'un média à l'autre n'est pas si fluide que cela. D'autre part, cela suppose que l'on puisse isoler des éléments hors contexte et les re-combiner sans qu'il y ait un travail d'écriture associé.

[DC-19] Ensuite, cela fait bon marché de toutes les difficultés que l'on rencontre lorsque ce passage se fait de manière automatique.

[DC-20] Enfin c'est faire l'impasse sur la rémanence des formes classiques, qui sont en permanence mimées dans les médias informatisés. Le rôle de l'impression dans la consultation des documents n'est pas à négliger et de nombreux sites, à commencer par les fournisseurs de courrier électroniques proposent désormais une « version imprimable » distincte de l'affichage à l'écran. « L'écrit d'imprimante, rappelle Emmanuel Souchier, présente ceci de particulier qu'il renoue avec le lien indéfectible que l'écriture entretient toujours avec son support. » [3]

[DC-21] Dans le même ordre d'idées le succès du format PDF devenu un véritable standard témoigne d'une volonté de ramener le document numérique dans des formes lisibles qui rappellent celles de l'écrit imprimé classique, tout en intégrant une dimension architextuelle présente dans l'appareillage de lecture qui encadre le document proprement dit.

Rôle de l'outil.

[DC-22] Avoir besoin d'un appareillage technique pour déchiffrer un document n'est pas une caractéristique qui suffit à elle seule à caractériser le document numérique puisque, comme nous l'avons vu, toute œuvre caractérisée par sa « reproductibilité technique » [4], nécessite, pour être abordée, un appareil reproducteur.

[DC-23] Cependant la situation d'un spectateur face à un écran de cinéma ou un téléviseur, n'est pas la même que celle du lecteur qui consulte un document sur une interface d'ordinateur ou sur son « assistant personne numérique ». Dans ce dernier cas, le document ne s'affiche pas seul, ou rarement, mais toujours dans le cadre ou

dans la mise en abîme de cadres qui jouent à la fois le rôle d'un appareil de lecture et d'un appareil d'écriture.

[DC-24] L'équipe qui a travaillé sur les médias informatisés pour une étude commandée par la Bibliothèque publique d'information, a qualifié d'architextes ces éléments par lesquels le texte parvient à l'existence dans le cadre des médias informatisés.

[DC-25] « Du banal traitement de texte au logiciel d'écriture multimédia, on ne peut produire un texte à l'écran sans outils d'écriture situés en amont. Ainsi le texte est-il placé en abîme dans une autre structure textuelle, un « architexte », qui le régit et lui permet d'exister. Nous nommons architexte (de arkhè, origine et commandement), les outils qui permettent l'existence de l'écrit à l'écran et qui, non contents de représenter la structure du texte, en commandent l'exécution et la réalisation. Autrement dit, le texte naît de l'architexte qui en balise l'écriture.

[DC-26] Structure hybride, héritée tout à la fois de l'informatique, de la logique et de la linguistique, l'architexte est un outil d'ingénierie textuelle qui jette un pont nécessaire entre la technique et les langages symboliques. » [5]

Pour reprendre une de nos interrogations du début, on voit ici le caractère fécond de l'interrogation inter-disciplinaire. L'analyse littéraire, à la suite par exemple de Genette (dont le concept d'architexte est différent de celui évoqué ici), a mis en lumière les notions de para et de péritexte. Ce qui, en dehors ou à côté du texte, en forme malgré tout partie et le déterminent.

[DC-27] Dans la documentation traditionnelle, il n'est pas question de confondre le document et son système de classification. La fiche est dans le meuble central et le livre est sur le rayon de la bibliothèque. L'une constitue le signe de l'autre et son référent. Mais dans le document numérique, des parties adjacentes ou sous-jacentes, qui peuvent être l'affichage de l'architexte du navigateur web, ou le descriptif des « méta-données » contenues dans le fichier source font partie intégrante du document, qui apparaît dès lors comme une agglomération d'éléments qui se présentaient auparavant dans des formes spécifiques et disjointes.

[DC-28] L'expression « document numérique » apparaît in fine, somme toute, mal formée, tout comme l'est celle des « nouvelles technologies de l'information » [6]. Elle est confuse, car tout en mettant au centre de la réflexion une dimension technique qui nous paraît trop souvent ignorée, elle uniformise une trajectoire du document qui se caractérise plus par une série de métamorphoses successives que par une trajectoire linéaire dans un environnement technique donné.

[DC-29] Tout document produit au début du 21^e siècle est, à un moment de son cycle, numérique. Mais dans ses métamorphoses successives, il a acquis des parures formelles qui le distinguent en fonction des supports sur lesquels il est destiné à être appréhendé. L'identifier de façon définitive par l'hypostasie d'un seul de ses moments

est courir le risque de décaler le champ de l'innovation là où il n'a pas de raison d'être et de perdre de vue les dimensions immanentes de la lecture qui resurgissent toujours par-delà les modifications de l'objet : le sens formel, l'ordre visuel, la dimension graphique, qui, par-delà les exigences nécessaires de la fabrication technique, caractériseront encore longtemps la réception de l'écrit.

[DC-C]

[1] Voir sur ces sujets, Philippe Breton, L'utopie de la communication, La Découverte, 1997 et Patrice Flichy, Une histoire de la communication moderne, espace public et vie privée, La découverte, 2004

[2] Carolyn Marvin, When old technologies were new, Oxford university press, 1990, voir aussi sur les mêmes thèmes, Patrice Flichy, Une histoire de la communication moderne, espace public et vie privée, La découverte,

[3] Emmanüel Souchier, Lorsque les écrits de réseaux cristallisent la mémoire des outils, des pratiques, Interdisciplines, the future of web publishing, Centre Jacques Cartier, Janvier 2003, <http://www.interdisciplines.org/desfispublicationweb/papers/18/1/5>

[4] Walter Benjamin,

[5] Emmanuël Souchier, Yves Jeanneret, Joëlle Le Marec, et al., Lire, écrire, récrire, Objets, signes et pratiques des médias informatisés, Bibliothèque publique d'information, Paris, 2003, p.23

[6] Yves Jeanneret, Y a-t-il (vraiment) des nouvelles technologies de l'information, Septentrion, 2001.

Un cycle panoptique pour l'analyse des documents

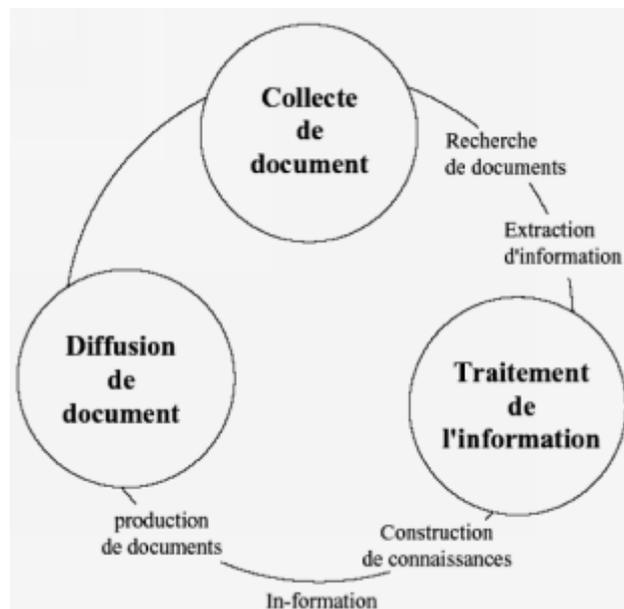
Gabriel Gallezot

[GG-1]

[GG-1] En complément ou en redondance du texte de Roger T Pédaque [7] qui définit le document numérique par les axes forme, signe, médium, ce texte, soumis à vos commentaires, livre un modèle heuristique d'analyse des documents, notamment ceux marqués du sceau « numérique ».

[GG-2] Le cycle panoptique veut embrasser sous forme d'un procès le document dans sa genèse et ainsi trouver dans ses racines des éléments de définition capable de l'appréhender dans son immatérialité*.

[GG-3] Nous inscrivons la constitution du document dans une double relation triadique. La première triade est composée des actions entreprises sur le document : la collecte, le traitement, la diffusion. La seconde est structurée par les opérations de transformation de l'information où la connaissance est la formation des idées, l'information est la mise en forme des connaissances (in-formation) et l'information inscrite sur un support constitue un document.



[GG-4] Les actions de l'ensemble « collecte/traitement/diffusion » de l'information sont réalisées par un binôme indissociable d'outils et de méthodes : les technologies

intellectuelles [4]. Celles-ci permettent des découvertes informationnelles qui se transformeront, après un processus créatif, en connaissance. Ainsi, nous distinguons l'Information retrieval (IR) de l'épistémé mais soulignons leur appartenance au même processus [3]

[GG-5] Dans notre proposition la notion intrinsèque de la forme importe peu, qu'il s'agisse d'un codex ou une d'une carte sémantique (fouille de texte) ces documents font partie d'un même processus. Néanmoins, la forme à un impact sur la capacité, la vitesse à produire des connaissances, sur la « motorisation » du cycle. La notion de forme est retenue pour sa fonction endogène au procès

[GG-6] Nous focalisons sur le signe, et plus spécifiquement sur un corpus de signes, indiquant plus encore que la forme importe peu dès lors que les unités informationnelles font sens auprès du lecteur (humain, machine, humain + machine).

[GG-7] Le médium, notion de support élargie par le concept de communication, ne recouvre pas uniquement les supports physiques sur lesquels sont inscrites les informations, mais aussi les éléments sociétaux par/sur lesquels l'information est véhiculée.

[GG-8] Avec Schamber [8] et Buckland [2] nous introduisons la définition d'un document numérique ainsi : un document numérique peut donc être défini comme une représentation numérique d'une preuve, qui doit pouvoir être réutilisée (reused) dans un autre processus de traitement. Une troisième notion essentielle à la définition d'un document concerne sa localisation, il devra renfermer l'information nécessaire à son repérage dans une collection. [5]

[GG-9] Trois concepts clés émergent alors. Ils offrent une autre façon d'envisager le découpage « forme, signe, médium ». Dans le contexte que nous avons dressé avec le cycle ils permettent d'appréhender les notions de : Granularité (signe), plasticité/versatilité (forme+signes), traçabilité (forme+signe+médium)

[GG-10] La granularité du document indique sa capacité à être séquencé, coupé, égrainé en unités informationnelles porteuses de sens. La plasticité concerne la capacité du document à être polymorphe, à recomposer des unités informationnelles selon des usages, des besoins, des situations différentes. La traçabilité répond à l'exigence de la preuve, essentielle à la notion de document [1].

[GG-11] Ces trois éléments indissociables et nodaux du document numérique mettent en évidence les besoins du traitement de l'information, non plus centrés sur la syntaxe mais sur le sens. Il ne s'agit plus, par exemple, de chercher des documents mais des morceaux de documents qui mis en forme selon des choix opérés par l'utilisateur proposeront de nouveaux documents tout en conservant le lien avec la source documentaire dont ils proviennent [6].

[GG-C]

Note :

* : Issu du cycle de l'IST, construction à partir des modélisation des systèmes d'information, cycle de vie du document, cycle de l'information .

Références :

1. BRIET, S. (1951). Qu'est-ce que la Documentation. Paris, Éditions Documentaires Industrielles et Techniques (ÉDIT), 47 p.
2. BUCKLAND, M.K. (1997). " What Is a Document? ". In : J. Am. Soc. Information Science , vol. 48, n°9, 1997, pp.804-809.
3. ERTZSCHEID Olivier, GALLEZOT Gabriel.. Chercher faux et trouver juste, Serendipité et recherche d'information, CIFSIC – Bucarest, 01 juillet 2003 : http://archivesic.ccsd.cnrs.fr/sic_00000689.html
4. FAYET-SCRIBE S. "histoire de la documentation en France : Culture science et technologie de l'information", CNRS éditions, 2000.
5. GALLEZOT Gabriel., Sansom F., Brunaud V., Gas S., Bessière P., "Normes et standards dans le processus de traitement du document numérique en biologie moléculaire ", Solaris, n°6, 2000. [En ligne]. <http://www.info.unicaen.fr/bnum/jelec/Solaris/d06/6gallezo.html> .
6. GALLEZOT Gabriel, Exploration informationnelle et construction des connaissances en génomique, Les Cahiers du numérique, Vol3-n°3, Hermes/Lavoisier, 2002
7. Roger T. Pédaque, Document : forme, signe et médium, les re-formulations du numérique. 08 juillet 2003, http://archivesic.ccsd.cnrs.fr/sic_00000511.html
8. SCHAMBER, L. (1996). "What is a Document ? Rethinking the Concept in Uneasy Times ". In : J. Am. Soc.Information Sciences , vol. 47, n°9, 1996, pp. 669-671.

La forme matérialisée, ou l'impossible contrat de lecture

Yves Jeanneret

[YJ-1]

[YJ-1] Ma contribution consiste à discuter une question que révèle le texte collectif signé « RTP ». Je trouve la situation intéressante, sur le plan discursif : une tentative d'écriture convergente dans sa valeur de provocation pour une écriture divergente. Dans son effort de formuler dans un vocabulaire commun des problématiques liées à des histoires différentes, RTP rend visibles des enjeux et des ambiguïtés. Je souhaite contribuer à expliciter ces enjeux.

[YJ-2] Dans sa dimension d'explicitation et de diplomatie, RTP distingue trois modes d'existence du document, comme objet-forme, comme signe-contenu et comme sociabilité-mémoire. En somme, le document support (Innis), le document information (Otlet) et le document monument (Foucault). Je ne conteste pas cette économie rhétorique ternaire, que nous avons d'une certaine façon adoptée dans le collectif *Lire, écrire, récrire* en sous-titrant « objets, signes et pratiques ». Le besoin de l'exposé justifie cette triple approche, qu'il ne faut pas mener trop loin en suggérant que les communautés s'organisent surtout autour de l'une des dimensions.

[YJ-3] Certains chercheurs, dont je fais partie, essaient surtout d'expliciter les relations entre les trois, de montrer que les formes sont matérialisées et que cette matérialité textuelle est l'un des lieux d'exercice d'un pouvoir culturel. Je pense qu'il est surtout essentiel de conserver à l'esprit ce à quoi pouvait servir cette typologie, une sorte de repérage du terrain, qui de fait croise en permanence les trois perspectives, notamment dans les parties consacrées aux questions vives. Et de se garder de faire ensuite de ce repérage un appareil de nomenclature organisationnelle de la recherche, parce que de la disjonction entre les trois dimensions surgirait à coup sûr la stérilisation de la recherche.

Contre la disjonction, d'abord inévitable, des objets, des signes et des pratiques il faut faire l'effort de comprendre comment les uns s'articulent aux autres.

[YJ-4] J'entrerai dans cette question par la notion de forme, qui me semble révélatrice des enjeux posés par une réflexion sur le « document numérique ». « Document numérique » avec des guillemets, puisque ce que la question de la forme montre d'abord, c'est que le document ne peut pas se définir seulement par le fait d'être numérique.

[YJ-5] Je ne crois pas à l'idée d'objet immatériel, et je soupçonne R. T. Pédaque de ne pas y croire non plus, puisqu'il réintroduit à juste titre, partout, une forme matérialisée au nom de laquelle, seule, s'établit la communication. Que la forme soit incarnée ou non me semble la grande bataille, le grand enjeu que soulève Roger dans

chacun de ses moments problématiques.

[YJ-6] Les échanges auxquels a donné lieu la rédaction de Pédaque portent la marque de cette bataille entre, d'un côté, une conception de la forme réduite à un formalisme logique, c'est-à-dire à une modélisation/ segmentation du texte et de l'autre, une conscience du pouvoir culturellement discriminant et structurant d'une forme incarnée dans la sémiotique sensible d'un percevoir organisé. Forme flotte donc tout au fil du texte entre formalisme (Fodor) et configuration (Chartier). Les deux ne sont pas conciliables.

[YJ-7] On peut aborder cette question d'un point de vue théorique, et je défendrai, évidemment, la conception incarnée des formes, parce que cette dernière permet de rendre visibles les pouvoirs d'organisation matérielle du texte (je reviens ci-dessous sur cette notion). Elle permet de ne pas renvoyer le « numérique » à une simple mise en accès, disposition, organisation d'un savoir plus ou moins insensible à son incarnation. Je trouve aussi intéressante, ou plus intéressante, la réflexion que suscite Pédaque dans son effort pour tendre des liens entre communautés, parce qu'il met en évidence, finalement, une tension.

[YJ-8] Le « document numérique » (cet oxymore de forme matérielle et de formalisme cybernétique) apparaît comme un champ de bataille entre des préférences sémiotiques, entre ceux qui essaient de le métamorphoser en une pure incarnation de modèles logiques (un texte constitué entièrement de son double abstrait) et un effort pour créer, sur les médias informatisés, de nouvelles formes écrites. Et ce que montre le texte collectif, c'est que ces deux conceptions contradictoires du texte ne s'affrontent pas seulement sur le terrain de la description mais sur celui de la configuration des objets et des pratiques.

[YJ-9] Je trouve que Pédaque a le grand mérite de montrer la force de cet enjeu, parce que les acteurs de la modélisation informatique ont en quelque sorte le pouvoir de nier ce qui fait l'épaisseur historique du document et en même temps se trouvent recréer – comme dans une schize entre formalisme logique et forme perceptive – un monde incarné, mais hanté par le rêve de la « device independence », qui est la négation de la forme par le formalisme : l'incarnation d'un modèle désincarné de la corporéité.

[YJ-10] A mon avis, nous devons tenir fermement une conception matérielle et sociale de la forme : une aptitude socialement construite à qualifier de signes des objets visuellement organisés et une inscription, dans les dispositifs, d'un ordre du texte. Certes, le travail de modélisation tend à s'autoriser d'une prétendue immatérialité du formel, mais il se traduit toujours par l'imposition matérielle de formes.

[YJ-11] C'est pourquoi je suis en désaccord – c'est mon seul désaccord profond – avec la définition du signe comme contenu et intentionnalité. Et avec la conception du texte qui accompagne cette définition, le terme « texte » équivalant plus ou moins à celui de « contenu intellectuel », de matière idéale qui resterait inchangée à travers la métamorphose des supports. Il me semble que le signe demande à la fois la

matérialité, la forme et la perspective sociale, il n'y a plus de signe si le signe est seulement un synonyme du contenu ou du sens. Le signe n'est pas le sens, mais l'objet à partir duquel il est possible de faire du sens.

[YJ-12] C'est pourquoi le texte ne peut pas être identifié simplement à un contenu, à une intention, à une pensée. Le texte est ce niveau indispensable de l'analyse qui permet de reconnaître une conception communicationnelle dans le « document » numérique (ce n'est pas pour moi un document, c'est-à-dire un objet matériel isolable, mais un texte, c'est-à-dire une configuration de formes reconnaissables inscrite dans une matérialité).

[YJ-13] L'approche que je propose ici de la notion de forme explique pourquoi je ne souscris pas à l'idée de « contrat de lecture ». Cette notion comporte deux grandes ambiguïtés. D'une part, on ne sait si elle concerne, comme chez Veron, le fait que toute production de communication porte l'inscription de rapports communicationnels virtuels ou si elle suppose, comme chez Charaudeau, l'hypothèse d'une convention partagée sous-jacente à la communication.

[YJ-14] D'autre part, on ne sait si on doit l'entendre au sens fort, comme paradigme conventionnaliste de la communication, ou de façon beaucoup plus figurée, comme équivalent d'une relation de communication plus libre. Le problème – et la commodité apparente – de la notion de contrat, est qu'elle semble pouvoir rassembler en un continuum les propriétés matérielles des objets et les propriétés sociales de leur interprétation. Mais c'est, à mon avis, au contraire ce qu'il faut à tout prix éviter de faire.

[YJ-15] Je préfère en rester aux textes, mise en œuvre dans une situation de communication de formes matérialisées en un objet, et au jeu complexe par lequel des sujets mobilisent des principes interprétatifs autour de ces textes. Il y a une forme-texte différente dans les médias informatisés (c'est-à-dire une forme qui imbrique de façon particulière la modélisation logique, l'inscription visuelle des formes et le jeu des interventions écrites) et cette forme, non homogène, lieu de pouvoirs et de tiraillements, fait l'objet d'un ensemble d'investissements, dont certains trouvent l'occasion de s'inscrire dans la textualité, tandis que d'autres relèvent de la manipulation et du détournement de ces productions textuelles.

[YJ-16] Plus concrètement, je pense qu'une distinction très forte entre forme et formalisme d'une part, un abandon de la notion fallacieuse de contrat d'autre part, feraient gagner l'analyse en clarté. Ces critiques adressées à R. T. Pédaque doivent beaucoup à la formulation qu'il donne des enjeux de la bataille documentaire (ou plutôt textuelle) contemporaine.

[YJ-C]

Le document constructif comme outil de la pensée.

Sylvie Leleu-Merviel

[SM-I]

[SM-1] Le texte de Roger T Pédaque [4] appréhende le document dans son passage au numérique suivant trois entrées : l'entrée « forme » privilégie l'objet matériel ou immatériel, l'entrée « signe » fait du document avant tout un porteur de sens, l'entrée « médium » le considère en tant que vecteur de communication.

[SM-2] Un premier texte-commentaire [3] a montré que le document numérique est le résultat d'un processus complexe composé de 5 niveaux de processus emboîtés (de 0 à 4), procédant respectivement d'une entrée signe pour le 1er et le 4ème, d'une entrée forme pour le 2nd, d'une entrée médium pour le 5ème, et d'une entrée à la fois forme et médium pour le 3ème. Il paraît dès lors très périlleux, ou du moins fortement réducteur, de séparer les entrées et de n'en prendre en compte qu'une à la fois.

[SM-3] Le présent commentaire vise à ouvrir la réflexion dans une autre voie. Le texte de Roger T Pédaque [4], sans l'affirmer pleinement, semble sous-entendre en maints endroits que le document est conçu comme un objet relativement stable, au moins dans ses contenus mais aussi dans sa forme.

Ce présupposé est légitime lorsqu'il est inscrit dans la filiation de l'archivistique, de la bibliothéconomie ou des techniques documentaires, pour lesquels le document est un donné préalable qu'il s'agit de gérer. Il s'avère moins pertinent lorsqu'on le libère de son héritage historique pour considérer le document numérique dans ses manifestations résolument inédites.

[SM-4] Nous nous intéressons notamment au « document de travail » ou « document en construction », encore dénommé « document martyr » par certains usagers ancrés plutôt dans l'informatique, « work in progress » pour d'autres adoptant une vision plus créative.

[SM-5] Nous convenons de l'appeler ici « document constructif », en référence à la controverse qui secoue la communauté des mathématiques et que l'on peut brièvement résumer par la citation de Jean-Pierre Changeux [1] : « *Nombre de mathématiciens, qu'on appelle constructivistes, pensent qu'un objet mathématique n'existe que dans la mesure où on peut le construire... Les critères d'acceptabilité des mathématiques constructives sont plus rigoureux que ceux des mathématiques non constructives* ». Par analogie, le document constructif est donc celui qui se construit au fur et à mesure que s'élabore une pensée, et qui a pour objet de la formaliser.

[SM-6] On peut noter pour commencer que le texte de Roger T Pédaque [4] disqualifie cette sorte de trace formelle en tant que document, dans la section document comme médium, qui pose : « *un document donne un statut à une information, à un signe matérialisé. Il est porté par un groupe social qui le suscite, le diffuse, le sauvegarde et l'utilise... Le statut du document s'acquerrait sous deux conditions : l'inscription doit dépasser la communication intime (entre quelques personnes privées) pour devenir légitime et la légitimité doit s'affranchir de l'éphémère* ».

[SM-7] Peut-on réfuter la socialité comme condition d'octroi d'un statut au document, en affirmant que le groupe social qui le légitime peut se limiter à un seul individu, à condition que la sociabilité, i.e. la potentialité à devenir un objet social, soit présente ? Ainsi, par exemple, le présent document que je suis en train d'élaborer dans l'ici et maintenant de l'échange interactif avec ma machine possède, en cet instant où j'écris, une portée sociale limitée à moi-même. Cependant il sera peut-être diffusé en ligne et soumis à la discussion, voire à la controverse d'une communauté dont je suis dans l'incapacité d'estimer la taille à ce jour. Ou peut-être ne le sera-t-il finalement pas. Est-il pertinent d'admettre que sa destinée détermine son statut ?

[SM-8] En outre, le présent texte souche sera tronçonné en petits paragraphes qui appelleront chacun une arborescence (que nous espérons foisonnante) de commentaires, et ceci pendant peut-être plusieurs dizaines d'années (pourquoi pas ?). Est-ce un document ? Si la réponse est oui, celui-ci acquiert alors deux propriétés nouvelles :

- [SM-9] Il est dynamique, c'est-à-dire indéfiniment variable dans le temps, sans jamais se stabiliser sur un régime permanent dans lequel il n'évoluerait plus après l'extinction d'un régime transitoire plus ou moins long. Ceci revient à acter la différence entre un document qui évolue parce qu'il n'est pas achevé et un document qui évolue indéfiniment parce que sa nature est fondamentalement évolutive, i.e. dynamique.
- [SM-10] Il est collaboratif et partageable, c'est-à-dire élaboré par interactions dynamiques entre membres d'un groupe social qui l'utilisent comme véhicule pour échanger.

[SM-11] Enfin, ce type de document a pour fonction principale d'accompagner, de seconder un processus créatif en cours d'élaboration : il est outil de la pensée, appareillage intellectuel, support projectif de l'avènement du nouveau.

[SM-12] Il convient de déterminer en quoi le changement de technique, notamment dans le passage de la plume chère au penseur « classique » (cette plume lui servant

aussi bien à produire du texte verbal, des schémas, des partitions, des figures ou des équations...) au clavier de l'ordinateur et ses multiples fonctions de modification, sont susceptibles d'influencer le processus créatif lui-même, de transformer la pensée dans ses enroulements et ses structures d'émergence. Ce questionnement se décompose en deux points de vue :

- [SM-13] Assistera-t-on à une transposition numérique des outils de créativité connus [2] ? A quand le « brain-storming » ou « remue-méninges » hypertextuel via Internet ? Cette question procède une fois de plus d'une mise en abyme, car n'est-ce pas précisément ce que la présente démarche tente de mettre en place ?
- [SM-14] Verra-t-on émerger de nouveaux outils de créativité totalement inédits, et si oui, lesquels ? Ces nouveaux outils seront-ils à même de faire naître des idées de forme ou de structure nouvelles ? Les innovations relatives au processus d'inscription sous forme de trace conduiront-elles à un renouvellement de la conduite de la pensée elle-même ?

[SM-C]

Références :

- 1. CHANGEUX Jean-Pierre, CONNES Alain. *Matière à pensée*. Paris, Editions Odile Jacob, 1992.
 2. DE BRABANDERE Luc. *Le management des idées. De la créativité à l'innovation*. Paris, Editions Dunod, 2002.
 3. LELEU-MERVIEL Sylvie, « Hyperdocument : appréhension topologique d'un espace imaginal sans clôture », en soumission dans *Information, Interaction, Intelligence*, 2004.
 4. PEDAUQUE Roger T., « Document : forme, signe et médium, les re-formulations du numérique ». http://archivesic.ccsd.cnrs.fr/sic_00000511.html, 08 juillet 2003.

Les commentaires

Niels W.L. , University of Tromso, Norway

[SM-2] In an article submitted to I3, I try to argue that any kind of document is a result of a process; just think of the fact in the English language, that the word Document is a verb as well as a noun / to document vs. a document ! If you go the Latin version, you can divide the word documentum and you get the verb doceo or docere, and last, but not least the suffix mentum. Mentum denotes acts or means and results of acts. If doceo means : I show, you can say that a document is a result of showing "something" by using some means/"outils" and if you have to tell somebody what that "something" is, you have to use some new means.

[SM-3] When you talk about document numerique or digital documents in contrast to former kinds of documents, you tend to talk about relatively stable paperbased documents forgetting the early history of document from the beginning of the 13th. century till around 1690, when a document was an oral phenomenon and thus an ephemeral object. One may view the stable document as an exception from the general ephemeral document ?

Totschnig Michael, UQAM

[DC-29] J'aimerais commenter un argument qui se trouve en conclusion du texte de Dominique Cotte qui rappelle que le document numérique est souvent seulement un moment dans la production d'un document papier. Il est vrai que ce passage entre l'écran et le papier est un phénomène important qui doit être pris en considération quand on analyse l'évolution des différents formats de l'écriture, et que la genèse d'un document écrit dans un contexte numérique mérite aussi une attention particulière comme le rappelle le texte de Sylvie Leleu-Merviel [SM].

Mais il me semble qu'aujourd'hui le terme de "document numérique" tire sa pertinence du fait que de plus en plus de documents ont un cycle de vie entièrement confiné aux dispositifs numériques, et que dans le cadre de ces dispositifs, le problème est justement l'absence de repères pragmatiques qui nous guident dans l'usage du support papier. Autrement dit, un document écrit nous informe sur son contexte d'énonciation grâce à un grand nombre de conventions qui se sont établis dans l'histoire de l'écriture. Ces

conventions concernent en grande partie ce que Genette appelle le paratexte, c'est-à-dire l'enveloppe qui assure l'interface entre lecteur et texte, sans laquelle celui-ci ne serait pas déchiffrable. Grâce à ce paratexte, nous savons attribuer un texte à un auteur, à un contexte institutionnel, à une date, à un domaine de référence, à un lieu géographique. Le paratexte nous permet de comprendre la fonction du document, sa pertinence pour nous. Pour le document numérique, l'organisation de l'espace paratextuel ne va pas de soi, et le défi de l'invention de nouveaux genres documentaires numériques, réside dans la normalisation de cet espace, l'émergence de nouvelles conventions qui régulent le statut social des documents.

Une autre manière d'aborder ce problème consiste à s'interroger sur le processus de l'énonciation d'un document. La raison pour laquelle l'impression sur papier d'un texte numérique peut être perçu comme un changement important de statut, c'est justement que dans beaucoup de contextes, l'énonciation d'un document nécessite l'inscription sur un objet persistant, et une certaine transaction au moyen de cet objet. Nous signons une lettre, la mettons dans une enveloppe adressé à une autre personne, et la confions à une boîte à lettre. Socialement, ce processus équivaut à un acte social, à une énonciation. Dans un contexte numérique, des processus équivalents se sont développés ou souvent restent encore à inventer. Le courrier électronique est un bon exemple. Qu'est-ce qui fait que l'acte de cliquer sur l'icône "Envoyer" dans un logiciel de messagerie change radicalement le statut d'un texte que nous étions en train de rédiger? La coopération d'un grand nombre d'agents informatiques distribués dans un réseau est aujourd'hui tellement stabilisé, que nous savons que ce simple acte initie une série d'interactions qui aboutissent à la présentation de ce même texte dans un autre cadrage sur l'interface d'un autre dispositif auquel le destinataire du message a un accès privilégié. C'est la conventionalisation de toute la procédure menant de l'acte de l'envoi à celui de la visualisation qui assure que le texte devient document sans qu'il soit nécessairement imprimé sur papier. La présentation paratextuelle de ce document inclut une série d'informations qui contextualisent le document, auteur, date d'envoi, sujet. Nous savons aussi que le format informatique de ce document peut contenir d'autres informations d'ordre plus technique qui ne s'affichent pas nécessairement dans l'interface de l'usage, mais qui peuvent être invoqués par des procédures spécialisées, et qui peuvent nous renseigner plus précisément sur le contexte de l'énonciation. Par exemple, l'en-tête d'un courriel peut spécifier les identifiants d'autres messages d'un même fil de discussion. Cet exemple simple aujourd'hui devenu trivial permet d'illustrer que la constitution d'un document numérique équivaut à la définition de son contexte d'énonciation

Zeller Jean-Daniel, Hôpitaux universitaires de Genève

[1]

[YJ-4] Non seulement le document ne peut pas se définir par le fait d'être numérique mais

même sous sa forme conventionnelle le concept est trop polisémique pour qu'il soit utilisable. J'ai tenté (dans un article à paraître cet automne dans la revue Documents numériques) d'utiliser les éléments définis dans le document Pédaque pour établir une typologie des documents qui permettra de faire avancer ce débat. Or les catégorisations couramment utilisées pour définir les données, définir les codes, ou définir les structures de document ne renvoient pas à des champs homogènes, clairement délimitables (à simple titre d'exemple: qui peut définir un documents "multimedia").

Mon intuition est que se sont les liens qui forment le sens du document et que ces liaisons ont jusqu'à présent été implicite. Le transfert vers le numérique nous force à les expliciter.

[YJ-13] Je suis aussi dubitatif quant à l'opérativité du concept de contrat de lecture, mais je ne le rejetterais pas pour autant sans examen. Là encore il faut expliciter les différentes natures de ce contrat. Pour moi je vois 3 macro-domaines qui demanderaient une analyse plus fine de la part des spécialistes:

- la lecture d'évocation, auquel appartient le domaine littéraire, le contrat repose alors sur le plaisir émotionnel reçu
- la lecture de réflexion, auquel appartient la production scientifique et dans laquelle on doit discuter de l'entrelacement entre le plan des données (faits) et le plan du document (arguments)
- la lecture transactionnelle, auquel appartient les documents d'affaires, les contrats de tous ordres, lesquels font en général appel à tout un appareil de méta-données, implicites ou explicites.

[2]

[DC-5] La perversion du terme multimédia illustre bien les glissements de sens non-contrôlé dans le champs qui nous occupe. Au départ, il s'agissait d'un « paquet » dans lequel des médias différents étaient physiquement rassemblés autour d'un thématique de sens/contenu commune (livre et disque, diapositives et bande sonore, etc.) et le terme était parfaitement approprié. L'apparition du codage numérique universel (l'alphabet, qui veut qu'un même octet puisse représenter un nombre, une lettre, une note, un pixel, ou une valeur logique) et d'un support normalisé (le CD-ROM) a permis de fusionner ces mêmes contenus sur un seul support, un « Unimedia », mais l'ancienne dénomination, installée dans l'usage, y est restée attachée.

[DC-12] En effet le document numérique n'est pas nouveau en tant que document, mais sa forme technique interroge les finalités du document dans la mesure où la non transparence de sa forme de codage, et la dissolution de son « unicité » fait douter de sa capacité à « documenter » (au sens de [SM-6] : donner un statut à une information). Un des exemple le plus frappant dans ce sens est l'existence de site web imitant des sites «

officiels », quels sont dans ce cas les éléments formels qui nous permettent de reconnaître la parodie de la tromperie ?

[[DC-15](#)] Le plagiat numérique contient son propre antidote, il suffit d'interroger un moteur de recherche avec les phrases suspectes pour en découvrir la source*encore faut-il savoir le faire et/ou y penser (il est clair que certains textes sont plus accessibles que d'autres).

[[DC-21](#)] Le succès de PDF présente un paradoxe. Il est assez clair que sa large diffusion en tant que format de communication vient de sa similitude avec l'aspect du document traditionnel. Mais même s'il permet certains verrouillages assurant sa fixité qui en fait un quasi-document (et qui seront probablement renforcés dans le projet de norme PDF/A, visant à produire des documents archivables à long terme), peu de producteurs utilisent pleinement ces fonctionnalités, ce qui permet de retranscrire des documents PDF dans d'autres formats pour les manipuler à nouveau.

[[DC-25](#)] Le concept d'architexte me semble prometteur, si on l'associe à d'autres concepts du champ. La piste de réflexion serait la suivante :

architexte structure fondatrice (forme au sens d'archétype) pourquoi ou pour quoi on écrit
texte contenu (données) ce qu'on écrit
contexte méta-données ce qui informe l'écrit. Dans ce cadre, l'hypertexte ne serait qu'un « texte au carré », mais où certains textes relient le texte au contexte. Il faut aussi remarquer que le contexte est sans fin, ce qui implique que l'architexte représenterait un « noyau dur » de sens, une volition, qui deviendrait de plus en plus fluide plus l'on suit les pistes du texte et du (des) contexte(s). (parenthèse : on rejoint une citation des pré-socratiques : le fond du fond, c'est la forme).

[[DC-27](#)] La jointure des contenus avec leur méta-données sont une nécessité technique dans la gestion des documents numériques « actifs ». Par contre le modèle d'archivage à long terme de ceux-ci, tel qu'il est décrit dans la norme OAIS, vise à nouveau à les disjointre dans des système de gestion à long terme (voir **)

[3]

[[SM-1](#)] La « forme » d'un objet immatériel est un contresens (à moins que l'on se place du côté des idées platoniciennes). Même virtuel un objet a toujours un support, la difficulté conceptuelle réside dans le fait qu'il n'a plus un support « unique » et présent comme dans le cas d'un document conventionnel, mais un support potentiellement multiple et distant, et que c'est peut-être justement ce fait qui pose problème dans le document numérique.

[[SM-3](#)] Tout à fait d'accord sur l'héritage historique qui nous fait présupposer qu'un document est stable dans le temps (alors que la numérisation bat en brèche ce

présupposé), il n'empêche que l'on garde aussi des « état antérieurs » dans ce monde là (les manuscrits d'*uvre éditée, les versions successives de projets de lois, etc.). Dans ce domaine la confusion conceptuelle existant entre les termes donnée-information-document obscurcis singulièrement la discussion. Pour éviter cet écueil et tenter de clarifier les notions, on peut tenter de verbaliser ces substantifs :

- Avoir des données c'est collecter des faits bruts (mais le sont-ils jamais ? voir « Nos pensées créent le monde », de Martine Castello et Vahé Zartarian, ed. Laffont 1994, accessible en ligne sur <http://creation.net/weid/txt-weid/sommaires.htm>) (note : le fait que l'on ne puisse pas dire « être donné » me semble significatif du hiatus conceptuel que je souligne)

- Etre informé c'est recevoir des données mises en forme (cette forme peut avoir l'état éphémère du journal télévisé) et cette forme implique fatalement un contexte sous-jacent, soit formalisé (la mise en page du journal ou de l'article), soit informel (la culture supposée partagée entre le metteur en forme et le récepteur, mais aussi la manipulation, qui réside entre autre précisément dans la mise « hors-contexte » de l'information).

- Etre documenté c'est avoir l'accès à des données « informées » en document. Il semble donc bien que c'est la « fixité » relative du document qui en constitue la valeur informative, et par conséquent « formatrice », on retrouve ici la source étymologique de docere.

Cependant, la culture documentaire occidentale a très vite pris l'habitude de « déconstruire » le document en informations ou en données pour en reconstruire un autre, de manière cyclique et récurrente (voir le texte de Gabriel Gallezot [GG]). Le numérique ne fait alors qu'amplifier et accélérer le phénomène de recomposition des idées et du savoir qui a fait la fierté des universités. Cette accélération de la dissolution/cristallisation des idées n'est cependant pas sans danger dans la mesure où sa rapidité engendre une certaine perte de contrôle. La publication (au sens large) du document comme une « photo » de la pensée ou de l'action, en permettant un arrêt sur image, permet ainsi le contrôle à posteriori, la validation (ou l'invalidation) des hypothèses ou des actes (on rappellera à ce propos qu'un des termes allemand pour document se dit Akten, ce qui correspond au concept de Record anglo-saxon, et rejoint une des définitions du terme acte, au sens de document officiel en français, ou pour le moins respectant une certaine forme, le plus souvent légale).

[SM-6 à SM-7]. Conséquence du commentaire en SM-3, cela signifierait qu'un objet numérique ne pourrait avoir le statut de document que dans la mesure où un certain nombre d'exigences de forme le rende comparable à un document de type conventionnel. Malheureusement, il n'existe pas de typologie des documents conventionnels uniformément reconnue, qui les classerait sur le plan du statut de l'information qu'ils contiennent, ce qui est d'ordre quasi juridique (la tentative la plus pertinente dans ce sens est celle développé par Parent R. et Boulet N. (coordonné par), Conseil du trésor -

Sous-secrétariat à l'inforoute gouvernementale et aux ressources informationnelles, Les composantes d'un document électronique, Rapport du Groupe de travail sur les méta-données et les structures logiques, Collection en ingénierie documentaire : 2, Janvier 1999, réalisé dans le cadre du Chantier en ingénierie documentaire, [<http://www.tresor.gouv.qc.ca/doc/word/collection-02.doc>]).

Les conditions posées par le document Pédauque sont bien entendu insuffisantes et peut être pas pertinentes. La difficulté réside dans le fait qu'il n'y a sûrement pas qu'un seul statut de document valide.

[SM-8] Le document comme « work in progress » tel celui que nous composons actuellement est peut-être justement un nouveau genre de document (la question du genre des hypertextes, qui est parallèle, a fait l'objet d'une session entière d'une conférence sous le titre « genre in digital documents ». Voir Proceedings of the 32nd Hawaii International Conference on System Sciences - 1999, accessible en ligne à l'adresse <http://computer.org/proceedings/hicss/0001/00012/0001toc.htm>. Le triplet de définition du genre comme « content, form, functionality » utilisé par certains auteurs se rapproche fortement des hypothèses du document Pédauque).

Il est cependant significatif que les modérateurs du débat actuel aient d'emblée annoncé que la totalité des contributions seraient conservées, alors même qu'un texte de synthèse sera produit. Il s'ouvre alors une nouvelle problématique, qui est celle du « temps » du document. La réflexion humaine, bien que toujours en mouvement, semble devoir éprouver la nécessité d'arrêter le temps à certain moment, pour se permettre une « vue d'ensemble » qui lui permet de se structurer, quitte à mettre en cause cette structure immédiatement après. La question serait alors de savoir quelle est la granularité temporelle optimale de tels documents/processus.

Le cycle de vie du document préoccupe depuis plusieurs années les archivistes australiens et néo-zélandais pour lesquels le document n'est finalement que la trace d'un processus plus vaste qui produit la mémoire institutionnelle dans son entier (voir Frank Upward, Structuring the Records Continuum Part One: "Post-custodial Principles and Properties", Archives and Manuscripts 24(2) (Nov 1996): 268-285; [<http://rcrg.dstc.edu.au/publications/recordscontinuum/fupp1.html>] ; Part Two: "Structuration Theory and Recordkeeping", Archives and Manuscripts 25(1) (May 1997): 10-35. [<http://rcrg.dstc.edu.au/publications/recordscontinuum/fupp2.html>]). Il définit ce continuum en cercles concentriques selon les séquences suivantes : Un acteur produit des actes qui laissent des traces qui constituent des documents (dimension personnelle de création), une unité produisant des activités dont l'évidence s'inscrit dans des records (dimension fonctionnelle de captation), une organisation assumant des fonctions dont la mémoire est assurée par des archives (dimension d'organisation), une institution (au sens de corps constitués)

répondent à des objectifs qui produisent une mémoire collective également conservée dans des archives (dimension institutionnelle de généralisation). A chacun de ces niveaux un même document peut prendre une valeur différente.

Broudoux Evelyne, IUT de La Roche-sur-Yon

[1]

La trace écrite électronique est-elle un constituant du document numérique ?

Peut-on encore distinguer le process d'écriture de l'acte de publication ?

Telles sont les questions posées entre le document comme preuve de G.G. [\[GG-8\]](#) et le document comme trace de S.M. [\[SM-14\]](#)

La fonction publication est-elle en train d'évoluer avec les outils d'écriture ?

Non, si l'on considère que le processus d'écriture exige la possibilité de ne pas être lu, sauf l'écriture performance, l'écriture de chat ou l'écriture wiki.

Oui, si l'on compare la trace fixée par le livre et celle du flux électronique. Alors que le livre fixe l'écrit pour des siècles, le flux électronique crée le concept de document à la demande, instantané, à la volée. Ce document flux reflète des apports collectifs, transitoires, dont l'objectif n'est pas la permanence, mais la fabrication en continu de la trace.

Ceci n'apparaît pas en contradiction avec la preuve, qui ne reste preuve que jusqu'à sa controverse.

[2]

[\[YJ-14\]](#) La forme matérialisée ou l'impossible contrat de lecture ?

Je plaide pour le contrat de lecture incluant la manipulation du support ! L'analphabète tient le livre à l'envers : la lecture sur support papier est liée à un acte engageant le corps. L'informatique rend les formes physiques de la lecture abstraites. L'idée du contrat de lecture est justement excellente à utiliser pour le document numérique parce qu'elle inclut à la fois le rapport "auteur-lecteur" et l'interaction possible par le lecteur avec le dispositif technique de lecture : c'est-à-dire l'outil.

Comme on sait que l'outil informatique impose déjà sa propre forme de lecture, il est évident que le lecteur aujourd'hui doit savoir à des degrés divers manipuler des formes abstraites de la lecture. Cela devrait être le premier rôle des interfaces graphiques que de soutenir ces manipulations.

Ajoutons à cela le fait que ce texte lu peut être à tout moment modifié. Un livre reste inerte, sauf à tourner ses pages. Un texte numérique est malléable, à portée du lecteur impliqué alors dans une lecture (Weissberg).

Le contrat est toujours double : il inclut les processus de compréhension du texte-image et ceux impliquant la compréhension du fonctionnement du support.

Petite digression au passage : tout ceci vaut aussi pour l'écriture... je me sens ici contrainte dans un "concept éditorial" dont je ne maîtrise pas les formes finales de publication, ce qui a pour effet de me "paralyser".

Augier Marc , CERAM

Je vous propose d'ajouter à [SM-14] une description des "wikis", qui sont sans doute un de ces nouveaux outils inattendus qui pourraient bouleverser notre relation au document.

Les Wiki(s) sont des sites Web animés par un dispositif qui permet à tout internaute de modifier le contenu de la page qu'il lit. Comme tout dispositif informatique, ce sont d'abord des informaticiens qui ont imaginé ce système, les mêmes que ceux qui ont mis au point le Web pour servir de passerelle d'échanges au plus grand nombre, pas ceux qui cherchent à en faire le plus grand supermarché planétaire.

Wiki viendrait du tahitien "wiki wiki" qui veut dire vite, parce que sur un wiki tout va vite, très vite. Une autre idée qui sous-tend les wiki nous vient du Zen: tout est inachevé, tout est éphémère comme le décrit Steve Wainstead sur le site du premier Wiki

[<http://c2.com/cgi-bin/wiki?WikiWikiWeb>]:

/Le tout est de rendre le formulaire EditText
<<http://webintec.ceram.fr/wikiwip/index.php/EditText>>
simple et la recherche ChercherUnePage
<<http://webintec.ceram.fr/wikiwip/index.php/ChercherUnePage>> rapide./

/En plus d'être rapide, ce site aspire aussi aux idéaux du Zen généralement réunis sous le nom de WabiSabi

<<http://webintec.ceram.fr/wikiwip/index.php/WabiSabi>>. Le Zen trouve la beauté dans l'imparfait et l'éphémère. Lorsque vous y parvenez, vous n'avez besoin de rien d'autre./

/L'Univers tend vers le chaos. L'Univers tend vers Wiki./

Pour revenir à notre sujet, les Wikis servent à fédérer des communautés de gens motivés par un même projet d'écriture. Chacun est libre de modifier tout ce qui ne lui plaît pas, d'ajouter tout ce qui lui semble bon d'ajouter, retirer tout ce qui lui semble superflu. Faisant cela, il sait que le prochain internaute qui lira sa page pourra en faire de même.

C'est donc la logique qui a commencé à être mise en place sur ce forum, poussée jusqu'à son extrême, les modifications n'étant plus sous le contrôle d'humains mais du site lui-même.

Comme pour [SM-8] les textes dans un Wiki sont tronçonnés en petits paragraphes, et disposent de liens hypertextuels pour les lier entre eux. Ces liens ne suivent pas une arborescence mais tissent une véritable toile, au fur et à mesure que des modifications s'ajoutent.

La socialité est le ciment de ce dispositif, ce qui pourrait venir donner des arguments pour ne pas la réfuter comme dans [SM-7].

En effet, si des individus qui ne se connaissent pas participent bénévolement à un wiki, c'est pour se "socialiser" dans une communauté qu'il jugent conforme à leurs idéaux. Nous pourrions citer l'initiative d'encyclopédie libre de droits créée et enrichie de la sorte sur <http://www.wikipedia.org> comme exemple le plus typique de cet état d'esprit. Mais, comme je l'ai noté précédemment, tout auteur sur Wikipédia sait qu'il sera jugé critiqué, et que ses erreurs seront mises en évidence et soulignée aux yeux de tous. La portée sociale d'un écrit sur un wiki est donc forcément ("par construction" comme diraient les ingénieurs des wikis) l'élément majeur qui a animé son auteur.

J'espère avoir contribué positivement à faire avancer votre réflexion, voire ouvert de nouvelles portes?

Je suis en train de démarrer un wiki dans l'école où je sévis : <http://webintec.ceram.fr/wikiwip/> . La mise en oeuvre est assez douloureuse, le manque de structure déroutant, expérience à suivre donc....

Réty Jean-Hugues , Université de Paris 8

[YJ-8] Cette opposition entre deux camps ne m'apparaît pas comme une fatalité.

Ces "nouvelles formes écrites" que cherchent à créer les uns ne pourront prétendre à une identité bien déterminée que si un formalisme sous-jacent, incarné dans un (ou des) format(s) numérique(s) particulier(s) et standardisé(s), constitue un substrat stable - comme l'est le livre imprimé pour la plupart des formes écrites - sur la matière duquel l'auteur décide de l'organisation et de l'apparence de signes. Ce que dit SGML - et toute sa descendance - n'est pas une antinomie formalisme cybernétique/forme matérielle, mais un dualisme forme logique/forme de présentation.

Inversement, une "pure incarnation de modèles logiques" ne peut justifier sa pertinence, trouver sa validation sociétale, que par une (des) forme(s) de présentation que reçoit l'oeil humain (ou tout autre organe sensoriel... je n'exclue pas ici le braille).

En somme, il ne s'agit pas finalement de "conceptions contradictoires", mais de deux

étages complémentaires d'une fusée.

Ce substrat nouveau de (potentielles) nouvelles formes écrites ne va bien sûr pas sans poser des questions et soulever des problèmes, comme le montre avec beaucoup de pertinence le document de Roger T. Pédaque. Parmi celles-ci se trouve la perte de contrôle - partielle mais réelle - de l'auteur (terme pris ici dans un sens très large) sur la forme de présentation effectivement donnée en réception au lecteur (toujours pris au sens large). Cette perte de contrôle se situe à deux niveaux.

Le premier niveau est celui du terminal de sortie. Il nous faut ici distinguer entre la forme de présentation effectivement donnée en réception au lecteur - que j'appellerai *forme de rendu* - et la spécification numérique d'une forme de présentation (par exemple : un fichier css) - que j'appellerai *spécification de présentation*. Pour une même forme logique et une même spécification de présentation, la forme de rendu différera sensiblement en fonction du terminal de réception, que celui-ci soit un écran à tube cathodique, un écran plat ou bien un vidéoprojecteur ; et la réception de ce rendu par un lecteur en sera affectée. (Il faut ajouter à cela la grande variabilité de la composante informatique des terminaux de sortie : une même page Web peut produire des rendus sensiblement différents selon que l'on travaille sur un PC ou sur un Mac, avec Netscape ou bien Internet Explorer, selon la taille de la fenêtre...) Ces différences dans la forme de rendu, dues aux technologies employées actuellement, se verront peut-être amoindries dans le futur - en terme de technologies, il est difficile de présager du futur - mais il est douteux qu'elles soient gommées complètement. Elle sont en tout cas pour le présent inhérentes au numérique.

La seconde perte de contrôle se situe au niveau de la possible pluralité de spécifications de représentation pour une même forme logique. La question qui se pose alors est celle de la définition du terme *document numérique*, ou plutôt de l'extension de ce terme au contexte des langages informatiques issus de SGML.

Certains outils informatiques permettant de manipuler des *documents* s'essaient à recréer dans le cadre du numérique un idéal de forme de rendu comme essence de la forme documentaire numérique (il s'agit d'un idéal, car le numérique rend de toutes façons très aléatoire la conservation de la pure forme de rendu). Ces outils, qui utilisent par exemple les formats .pdf ou .doc, sont plus ou moins directement dérivés du monde du document papier (bien que la notion classique de document ne se limite pas au support papier), et sont d'ailleurs très largement orientés pour répondre à des besoins d'impression. Les conséquences sur la problématique du document de l'utilisation de ces outils sont en elles-mêmes difficiles à appréhender et méritent largement d'être discutées comme elle le sont dans le cadre du RTP-DOC.

La mouvance SGML introduit cependant une nouveauté essentielle : la séparation

effective de deux entités signifiantes^{1*} qui, bien qu'ayant été pensées depuis longtemps dans leur dualité (graphème - sémène), ne s'étaient jamais trouvées, au fil de l'évolution des dispositifs matériels, séparables physiquement. (A cela s'ajoute une seconde nouveauté essentielle qui est l'objectivation de la structure logique du texte, puis d'autres nouveautés encore avec la généralisation et la standardisation des métadonnées...) Parler de "document numérique" nécessite de déterminer où se situe le *document* dans cette séparation.

La réponse est loin d'être évidente. Yves Jeanneret semble considérer que le document ne peut être que le tout ("cet oxymore de forme matérielle et de formalisme cybernétique"), et de ce point de vue découle la contradiction énoncée.

1*. Il s'agit en fait d'une séparation en 2+1, la forme de présentation ne pouvant se comprendre qu'en distinguant spécification de présentation et forme de rendu. On pourrait de manière similaire parler de séparation 1+1 - format informatique/forme de rendu - dans le cas des formats .doc, .pdf ou d'images numériques.

Ertzscheid Olivier , Université de Toulouse

[1]

[\[SM_8\]](#)

Ce "texte souche", tour à tour "tronçonné, enrichi de "commentaires" pouvant eux-même s'étaler sur une temporalité indéfinie, ce texte-souche donc dont parle [SM] demeure à mon sens un document à la seule condition de penser ce document dans le cadre de ces instanciations successives, c'est à dire dans une logique de "versioning".

Ce problème du « versioning » désigne l'ensemble des manières de gérer, à l'échelle de l'hypertexte planétaire, les procédures permettant de rattacher un texte à un auteur (ou à un collectif d'auteurs), tout en permettant à chacun de s'appropriier - de se ré-appropriier - tout ou partie de documents produits par d'autres ou par eux-mêmes afin, premièrement, de limiter la prolifération « bruyante » des versions différentes d'une même information sur le réseau et deuxièmement, d'identifier la nature et les origines de ces modifications dans l'optique d'une gestion cohérente de l'ensemble des documents électroniques actuellement disponibles, indépendamment de leur format, de leur statut et en dehors de tout institution centralisée. Autant dire que plus qu'une problématique, il s'agit là d'un véritable « idéal ».

La première des réalités que recouvre ce problème est celle du travail coopératif : tout changement de version témoigne d'une activité coopérative avec d'autres ou avec soi-même, dans le cadre d'une session, elle-même différente de celle correspondant à la version antérieure du document en question. La question du versioning pose en effet celle du cadre temporel propre à la coopération dans un environnement hypertextuel, c'est-à-dire dans un temps qui est celui de la session.

En plus du travail coopératif, quatre champs de recherche sont directement liés au problème du versioning : l'ingénierie logicielle, la gestion de document, les aspects juridiques et l'archivage.

Le premier problème d'ordre pratique que pose la gestion du « versioning » indépendamment de l'échelle à laquelle il s'applique, est celui des méthodologies permettant de suivre l'évolution de ces différentes versions d'un document via une gestion de l'affichage dans des systèmes ou des environnements différents.

Trois tendances se dégagent :

- la première consiste à lier le document initial de manière référentielle à ses versions successives. C'est l'option la plus fréquente dans les systèmes existants ;
- la deuxième permet d'inclure les versions dans le document initial ;
- la troisième crée un nouveau document (un nouvel objet) à chaque nouvelle version, lequel reste lié aux versions précédentes par de simples fonctionnalités de butinage semblables à celles présentes dans les navigateurs. C'est la solution retenue dans le système Xanadu de Ted Nelson.

Enfin, le versioning ouvre à l'échelle du web des problèmes à la portée « philosophique » par ailleurs posés dès la construction des premières bibliothèques : il s'agit du rapport de causalité qu'entretiennent nécessairement les choix de conservation et de stockage de documents (numériques) et la manière dont ils influent sur notre rapport individuel et collectif à la connaissance.

[2]

[[SM-7](#)]

En complément du texte RTP-DOC (section : document comme médium), repris par [[SM-6](#) et [SM-7](#)] je propose d'interroger le document non comme "simple" "trace formelle" mais plutôt comme un authentique lieu technologique.

Dans l'ensemble des technologies d'écriture et de production de documents mobilisées

par l'hypertexte, c'est d'abord l'ancrage en un lieu qui est privilégié, qu'il s'agisse du lieu-adresse auquel le document se rattache, des lieux-récit qu'il organise, ou du lieu-écran sur lequel il s'affiche. Et cette tendance ne fait que se confirmer au gré des évolutions du web :

« Si les premiers documents HTML comportaient avec eux leurs ancres et constituaient un document 'auto-suffisant', la multiplication des appels de scripts (Imagemap, formulaires, génération automatique de pages W3 en fonction du contexte ...) tend au contraire à localiser fortement un document. » [Le Crosnier, 95]

Dernière en date, dans le cadre du web sémantique, la mise au point d'un nouveau protocole d'adressage (URI) appelé à remplacer les URL existants, et qui permet d'attribuer à chaque ressource, à chaque texte produit, un lieu unique et déterminé une fois pour toutes .

En plus et en parallèle de ces procédures d'adressage, hypertextes et documents numériques disposent d'une palette de composants hors-langue qu'ils font entrer dans le cœur des processus d'écriture en rendant possible un « cinétisme de l'inscription » :

« Le 'déjà et toujours là' de l'œuvre sur le support laisse place à une surface qui est en gestation. Au lieu d'être présent, le texte arrive, surgit, apparaît. Son apparition peut-être globale et instantanée, ou bien peut se présenter sous la forme d'un enchaînement d'arrivées contrôlées de ses éléments juxtaposés ou non. Le contrôle s'exerce aussi bien sur les intervalles de temps entre l'arrivée de ces éléments que sur l'emplacement de l'inscription, ainsi que sur la ou les directions de son déploiement. » [Papp, 95].

Chacune de ces temporalités de surgissement et d'affichage est ainsi doublement localisée : par l'adressage auquel elle renvoie et par l'espace écran qui lui est dévolu.

Avec la temporalité, viennent ces deux autres constituants que sont la topographie et le mouvement .

« Le texte est immuable sur le papier ; par contre, sur l'écran il a une durée, des temps forts, des temps faibles. D'abord il y a le temps de l'attente, puis le temps (le moment) de l'apparition, ensuite le temps de présence qui est la somme des temps d'états statiques, cinétiques, scintillants ou non-scintillants, latents ou réels du texte, ensuite vient le temps (l'instant) de la disparition qui est généralement suivi par un temps d'écho. » [Papp, 95]

L'étymologie de l'inscription nous renvoie à une écriture du dedans (in-scription), une écriture qui prend naissance à l'intérieur d'un objet (le texte) et d'une subjectivité (l'auteur). L'écran, surface non-inscriptible directement, n'est que la médiation, le reflet, la

fenêtre ouverte sur une inscription décalée, en-deçà, rapportée, importée d'un autre lieu. Seule certitude qu'il autorise : le lieu de l'inscription du document numérique n'est pas l'écran. Il est ailleurs. Alors peut-on parler d'une aporie du lieu rendant la notion même document utopique ?

« La technologie électronique ôte ou soustrait du texte l'écrivain et le lecteur. (...) Il y a tellement de niveaux auxquels ils sont déferés que le lecteur ou l'écrivain éprouvent de la difficulté à identifier le lieu du texte : est-il sur l'écran, dans la mémoire de l'ordinateur ou sur le disque dur ? » Bolter D.J. Cité par [Landow, 92]

L'écran, l'archive de l'ensemble des textes disponibles, l'adresse (URI) à laquelle ils sont accédés, la mémoire (locale) de l'ordinateur sur lequel ils sont parcourus, le support physique qui les contient (cédérom, disquette), telle ou telle des versions qui les composent ... la question du lieu du document, quel que soit l'aspect que l'on choisisse de retenir, semble constamment se dérober à l'analyse critique. Une des possibilités alors souvent choisie est de considérer le document dans son actualité, dans son actualisation, c'est-à-dire à un moment donné de lecture, d'activation et d'état d'organisation interne des unités qui le composent (dans le cas des générateurs de texte par exemple). Mais là encore, si nous envisageons l'un ou l'autre de ces trois moments, chacun peut être caractérisé par au moins deux des paramètres cités plus haut.

« (...) ubiquitaire habitant du cyberspace, l'hypertexte contribue à produire ici et là des événements d'actualisation textuelle, de navigation et de lecture. Seuls ces événements sont véritablement situés. Quoiqu'il nécessite des supports physiques lourds pour subsister et s'actualiser, l'impondérable hypertexte n'a pas de lieu. » [Lévy, 88]

Le lieu de l'hypertexte est l'envers de chacun de ses endroits. Il est un lieu abîmé, il s'inscrit dans un espace en permanence renouvelé de significations. C'est cet espace qui est adressable, qui peut être cartographié. C'est dans ces cartographies, dans cette approche topologique de l'espace qu'apparaissent puis disparaissent des foyers d'émergence de la textualité, dans le temps de la perception cognitive d'une individualité de lecture ou d'une collectivité d'écriture.

Le document numérique nous semble avant tout être une topo-GRAPHIE autour de laquelle se joue la renégociation constante entre l'espace d'une carte et celui d'un territoire.

[3]

[[GG-I](#)]

En sus, ou plus exactement en complément, du cycle panoptique proposé par [GG] pour l'analyse des documents, il me semble que la collaboration STIC-SIC sur le document devrait s'efforcer de penser celui-ci dans le contexte de son déploiement, c'est à dire dans l'hypertexte planétaire que viennent structurer ou dé-structurer ces documents numériques.

Quatre entrées permettent alors de problématiser différemment le "document numérique".

1. L'organisation hypertextuelle des documents comme médiasphère.

« On ne peut séparer une opération de pensée, (...) des conditions techniques d'inscription, de transmission et de stockage qui la rendent possible (...). Le système dominant de conservation des traces (saisie, stockage, circulation) sert de noyau organisateur à la médiasphère d'une époque donnée dans une société donnée. (...) L'évolution technique des moyens de transmission matérielle donne un fil directeur à la succession historique, apparition et extinction, des systèmes symboliques vivants pour tel ou tel état du monde. » [Debray, 91]

L'organisation hypertextuelle des documents numériques prend place de fait, dans le cadre d'une médiasphère : non seulement elle implique d'un côté des agencements – collectifs ou non – d'énonciation (individus) et de l'autre des documents (textes) mais elle permet également d'organiser « (...) un espace/temps particulier, c'est-à-dire qu'elle se caractérise par un régime de vitesses techniquement déterminé mais intellectuellement et socialement déterminant (...) » [Debray, 91].

Dans la topographie particulière de l'hypertexte, la portée de l'axiomatique médiologique selon laquelle « Mémorisation des traces et locomotion des hommes ne sont pas vraiment séparables. » est augmentée de manière significative .

2. L'organisation hypertextuelle des documents comme logosphère.

A l'analyse des instances d'énonciation présentes sur les réseaux et au vu des modalités de discours qu'elles produisent et s'approprient, se donnent à lire de nouvelles organisations mémorielles. Ainsi, derrière ce qui apparaît comme une problématique de l'énonciation dans un environnement hypertextuel, une double perspective se dégage : il s'agit de construire un discours , ce qui ne peut se faire qu'à la condition de connaître et de maîtriser les nouveaux codes d'une énonciation de type « piétonnière » [Clément 95] dans laquelle la parole se déplace et change constamment d'axe.

« Rien n'est plus essentiel à une société que le classement de ses langages. Changer ce

classement, déplacer la parole, c'est faire une révolution. (...) Or depuis près de cent ans, depuis Mallarmé sans doute, un remaniement important des lieux de notre littérature est en cours : ce qui s'échange, se pénètre et s'unifie, c'est la double fonction, poétique et critique de l'écriture. » [Barthes, 66]

C'est cette double fonction, « poétique » (énonciation piétonnière) et « critique » (construction d'un discours) qui peut définir la plupart des contours émergents d'une logosphère, et à travers elle, de répondre à la question du « mode de constitution de cet hypertexte planétaire ».

3. L'organisation hypertextuelle des documents comme mécanosphère.

La première particularité de l'organisation hypertextuelle n'est pas tant que les textes et documents y figurant soient reliés (ce qui était déjà possible dans d'autres systèmes bibliographiques ou de classement, bases de données, etc.) mais que ces liens constituent une entité à part entière conditionnant – c'est-à-dire posant et résolvant dans le même temps – les problèmes liés à la navigation dans cette masse de documents.

« Nous trouvons ici un concept, celui de « path-dependence » (dépendance par rapport au chemin), (...) Ces systèmes [reposant sur ce concept] présentent en général une grande multiplicité d'états asymptotiques, ou attracteurs, ou « structures émergentes ». » [Bougnoux, 93]

Les relations existant entre ces entités liens et les systèmes et types d'organisation dont ils permettent de rendre compte ont souvent été analysées. Il existe de nombreuses applications, de nombreux environnements de création, d'écriture ou de navigation hypertextuels, qui peuvent tous être rassemblés sous le terme de mécanosphère, soit parce qu'ils sont des machines abstraites, soit parce qu'ils permettent d'agir sur des niveaux d'abstraction de plus en plus complexes, soit parce leur utilisation affecte tout ou partie des propriétés d'abstraction du monde alors institué :

« (...) toute machine abstraite renvoie à d'autres machines abstraites : non seulement parce qu'elles sont inséparablement politiques, économiques, scientifiques, artistiques, écologiques, cosmiques – perceptives, affectives, actives, pensantes, physiques et sémiotiques – mais parce qu'elles entrecroisent leurs types différents autant que leur exercice concurrent. Mécanosphère. » [Deleuze & Guattari, 80]

Indépendamment du niveau de constitution de ces systèmes, de leurs champs d'application ou de leur identité machinique :

« (...) les machines technologiques d'information et de communication (de l'informatique à la robotique en passant par les médias) opèrent au cœur de la subjectivité humaine, non seulement au sein de ses mémoires, de son intelligence, mais aussi de sa sensibilité, de ses affects et de ses fantasmes inconscients (...) » [Mattelard, 97]

Ces liens, ces entités reliées et les différents types de navigation concernés répondent à la question de la production, de la transformation et du transport des discours structurant la logosphère.

4. L'organisation hypertextuelle des documents comme toposphère.

Quand les agencements d'énonciation présents dans la logosphère investissent et s'approprient les processus de liaison-navigation permis par les systèmes relevant d'une mécanosphère, certaines formes, certaines structures émergent de cette conjonction et façonnent l'ensemble de l'organisation. Ces formes sont principalement repérables, identifiables et analysables d'un point de vue topologique.

[4]

[\[DC-8\]](#)

D'autant plus "loin d'être définie" que la conjonction de ces deux termes (document ET numérique) soulève également un certain nombre de questionnements esthétiques : s'il est des documents numériques, à quelles conditions et à partir de quel moment peut-on alors parler "d'oeuvres" numériques ?

Techniques, opérations, machines ... tels sont les espaces ouverts à l'orée du XXIème siècle ; espaces omniprésents qui tendent à occuper une place majeure et incontournable dans notre saisie du réel, dans notre environnement cognitif ; espaces où se déploie le numérique sous toutes ses formes : hypertexte numérique, images numériques, réseaux numériques, supports numériques, mémoires numériques, matière numérique ... et bien sur "document numérique".

Or qu'est-ce en fait que le numérique ? " Ce qui appartient au nombres " nous indique notre dictionnaire, lui aussi numérique. Ainsi donc nos mémoires, nos textes, nos images appartiendraient aux nombres ; tout ne serait que chiffrage ...

" Chiffre : caractère dont on se sert pour représenter les nombres. "

" Nombre : se dit de l'unité et d'une collection soit d'unités, soit de parties de l'unité. "

L'informatique nous a peu à peu accoutumés à l'idée de ces " bits " sous forme de 0 et de 1 qui suffisent à rendre perceptible, lisible, acceptable et parfois même émouvante une réalité non plus numérique mais virtuelle. Nous savons que notre morceau de musique favori – ce sublime Impromptu de Schubert – qui passe actuellement sur notre chaîne n'est lui aussi qu'une suite de 0 et de 1. Nous en avons connaissance, certes, mais ce qui prédomine, c'est l'idée persistante de cette salle de concert où nous l'avons entendu pour la première fois et où toute idée de nombre était absente ...

On peut de plein droit s'interroger sur la légitimité de ce qualificatif de "numérique" appliqué au document dans la mesure où aussi loin que l'on remonte dans l'histoire des sciences et de l'humanité, le numérique est toujours fortement présent, et ses implications dans les sphères techniques et philosophiques voire politiques sont fondamentales et fondatrices.

Compter, c'est à dire exprimer une suite de nombres, a certainement été la première manifestation de l'intelligence humaine aux temps préhistoriques ; les premiers travaux purement spéculatifs de la pensée de l'homme furent certainement des recherches sur les propriétés des nombres ; les premiers " savants " furent des arithméticiens, comme le montrent les premiers écrits qu'ils ont laissés. Pythagore, Platon, Aristote, tous se sont intéressés à la théorie des nombres. Le mythique nombre d'or ou section dorée (qui correspond au partage le plus harmonieux d'une grandeur en deux parties) fait partie intégrante de l'histoire des arts. Bien avant que les textes ne soient numérisés, il n'est jusqu'au texte fondateur de notre civilisation judéo-chrétienne qui ne contienne son " Livre des Nombres ", quatrième livre du Pentateuque.

Ainsi donc pourquoi se restreindre à l'entrée dans l'ère du "document numérique" quand se dessine une humanité numérique, une évolution numérique ? Ce qui change fondamentalement, tel un redéploiement de l'immémorial paradigme biblique, c'est l'incarnation. " Au commencement était le verbe ... " et le verbe se fit chair. A son tour le nombre semble s'incarner, prendre chair et forme, pour devenir matière. Et s'il est avéré que c'est le changement de la " forme " qui fait les grandes révolutions – que l'on songe au passage du Volumen au Codex – c'est l'avènement de cette matière numérique qui, en parallèle aux questionnements des SIC et STIC, devrait également inaugurer un questionnement de nature plus "esthétique". En voici succinctement quelques raisons.

Les enjeux du numérique pour ce qui est du matériau hypertextuel, concourent à l'émergence de nouvelles problématiques qui défient les catégories esthétiques habituelles et qui, plus fondamentalement, remettent en cause la notion archétypale d'œuvre. Parmi la foule des nouvelles entités créées et des habitus cognitifs associés, on peut lister comme étant les plus caractéristiques et les plus pertinents :

- navigation et non plus lecture,
- nomadisme textuel,
- mutation des supports ...
- arrivée de ces “ nouveaux acteurs ” de l'énonciation et des changements qu'ils occasionnent dans la mise en place du paradigme discursif et dialogique,
- problématique de l'identité textuelle qui est – tout ou partie – fonction des modes de génération choisis,
- le positionnement de la critique face à ces nouveaux biais de l'analyse que propose l'utilisation des nouvelles technologies.

Il s'agit là de quelques uns des éléments qui font vaciller sur ses bases la notion d'œuvre, vacillement que Foucault avait déjà perçu :

“ Mais surtout les unités qu'il faut mettre en suspens sont celles qui s'imposent de la façon la plus immédiate : celles du livre et de l'oeuvre. (...) Quant à l'oeuvre, les problèmes qu'elle soulève sont plus difficiles encore. En apparence pourtant, quoi de plus simple ? Une somme de textes qui peuvent être dénotés par le signe d'un nom propre. Or cette dénotation n'est pas une fonction homogène : le nom d'un auteur dénote-t-il de la même façon un texte qu'il a lui-même publié sous son nom, un texte qu'il a présenté sous un pseudonyme, un autre qu'on aura retrouvé après sa mort à l'état d'ébauche (...) ? Et quel statut donner aux lettres, aux notes, aux conversations rapportées, aux propos transcrits par les auditeurs, bref à cet immense fourmillement de traces verbales qu'un individu laisse autour de lui au moment de mourir, et qui parlent dans un entrecroisement indéfini tant de langages différents ? (...) On admet qu'il doit y avoir un niveau aussi profond qu'il est nécessaire de l'imaginer auquel l'oeuvre se révèle, en tous ses fragments, même les plus minuscules et les plus inessentiels, comme l'expression de la pensée, ou de l'expérience, ou de l'imagination, ou de l'inconscient de l'auteur, ou encore des déterminations historiques dans lesquelles il était pris. Mais on voit aussitôt qu'une pareille unité, loin d'être donnée immédiatement, est constituée par une opération ; que cette opération est interprétative (puisqu'elle déchiffre, dans le texte, la transcription de quelque chose qu'il cache et qu'il manifeste à la fois). (...) l'oeuvre ne peut être considérée ni comme unité immédiate, ni comme une unité certaine, ni comme une unité homogène. ” (Foucault, 1987)

Partageant avec l'œuvre – mais à une échelle différente – la notion d'unicité, on retrouve le fragment dont parle Foucault et qui est un caractère essentiel (au sens étymologique du terme) du document numérique :

“ L'hypertexte [le document numérique] partage avec ces derniers [récits arborescents et pure combinatoire] la notion d'unités narratives fragmentées. Mais ses fragments ne sont

ni totalement structurés, comme dans les récits arborescents, ni totalement inorganisés comme dans des textes à combinatoire totale. L'hypertexte est donc une collection de fragments textuels semi-organisée. ” (Vuillemin & Lenoble, 1995)

Toute la difficulté consiste – et c'est là un autre des enjeux majeurs d'un questionnement esthétique à poursuivre – à faire varier cette échelle de l'unicité en y introduisant comme paramètres la notion de discursivité, l'idée d'énonciation, etc ... pour voir quel est le seuil d'analyse au-delà ou en-deçà duquel le fragment acquiert le statut d'œuvre, et jusqu'à quel point et selon quelles modalités une œuvre numérique peut être fragmentée sans se défaire du principe d'unicité qui la fonde.
