

Document trace et document source. La technologie numérique change-t-elle la notion de document ?

Marie-Anne Chabin

► **To cite this version:**

Marie-Anne Chabin. Document trace et document source. La technologie numérique change-t-elle la notion de document?. Revue I3 - Information Interaction Intelligence, Cépaduès, 2004, 4 (1). sic_00001020

HAL Id: sic_00001020

https://archivesic.ccsd.cnrs.fr/sic_00001020

Submitted on 5 Jul 2004

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

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

Document trace et document source. La technologie numérique change-t-elle la notion de document ?

Marie-Anne Chabin

archiviste-paléographe, cabinet d'expertise Archive 17, marie-anne.chabin@archive17.fr

Résumé

Le mot « document » s'emploie ordinairement pour désigner deux choses : d'un côté la trace volontaire qui résulte d'une activité humaine d'administration ou de communication, matérialisée grâce à l'écriture ; de l'autre, une source d'information où l'on peut puiser des connaissances. A l'ère numérique, en dépit d'une redéfinition radicale de la notion de support qui passe d'un morceau de matière à une chaîne matérielle et logicielle, le document garde cette double fonction d'enregistrement des faits ou du discours, et d'offre au questionnement du lecteur, pour peu toutefois que les contraintes du numérique que sont la fixation de l'information et la maintenance des objets d'information soient respectées.

Mots-clés : *document, document numérique, auteur, utilisateur, trace, source, techniques d'écriture, support.*

Abstract

The French word « document » is commonly used in two ways : first as the trace produced deliberately by a human activity, written down and materialized; second, as a source of information, of potential knowledge. In the digital age, the medium is becoming a complete hardware and software range and need a new definition. But the documents still have the double function of recording transactions and ideas, and offering answers to the readers' questions. However, the technology has to ensure that the content is fixed and reliable, and the information objects are maintained and preserved.

Keywords : document, digital document, creator, user, trace, source, writing techniques, medium.

1 INTRODUCTION

Cet article a été suscité par une phrase particulière de l'appel à contribution qui, synthétisant le texte de Pédaque et ses nombreux commentaires, envisage que, finalement, dans l'environnement numérique « un document n'aurait d'existence à proprement parler qu'à deux moments: éventuellement celui de sa conception par son auteur ... et surtout celui de sa re-construction par un lecteur ».

Cette formulation provoque la question de savoir si cette double existence du document, à supposer qu'elle se confirme, est une nouveauté due au numérique ou si elle a des origines plus profondes.

Les pages qui suivent tentent de préciser la notion de document au cours des siècles, tant du point de vue de l'auteur que du point de vue du lecteur et de voir si la révolution technologique de la fin du XXe siècle modifie cette notion duale de document et, si oui, dans quelles mesures.

2 IL ETAIT UNE FOIS DES DOCUMENTS

Si l'on considère le document dans l'acception très large et communément admise d'un support comportant une information lisible par l'homme, on se place du point de vue de l'utilisateur car cette définition ne précise pas la notion d'auteur. On est donc renvoyé à la question : « Qu'est-ce qu'une information lisible par l'homme ? » pour laquelle nous proposerons trois réponses. Une information lisible par l'homme peut être :

- un produit de l'écriture, au sens d'un système de notation ou de code du discours humain, autrement dit un texte ; ces documents-là remontent à l'invention de l'écriture, à 5000 ans d'ici ;

- une pensée humaine, exprimée sans le secours de l'écrit codé et donc directement « lisible » par tout humain : dessin, croquis, image ;

- un élément de matière, étranger à l'homme mais que celui-ci, grâce à ses connaissances et à facultés créatives, peut analyser, commenter, interpréter.

Les trois cas de figures ne sont pas étanches et on les trouve fréquemment mêlés : échantillon de matière commenté par un

scientifique, photographie illustrant la visite d'un bâtiment, etc., mais le plus souvent d'une façon qui donne à l'écriture la primauté de l'expression. De sorte qu'on peut légitimement rattacher au premier cas tout ensemble composite englobé par un discours écrit qui lui donne sens.

2.1 Le document-trace

Une conjonction activité-outil

Pour en revenir aux supports, on peut observer au cours de ce laps de temps et plus encore au cours des derniers siècles que le support et la forme des documents sont conditionnés par les outils et techniques d'écriture à la disposition de leur auteur au moment où il écrit. Le temps, l'époque de création du document a une influence évidente sur son existence physique.

Un document peut être décrit comme la conjonction entre une activité humaine et une technique d'écriture, elle-même subdivisée entre un code, ou plusieurs codes, pour transcrire ce qu'on veut exprimer, un ou plusieurs outils et une ou plusieurs matières pour matérialiser l'écrit. La conjonction de l'acte de compter les troupeaux avec la pratique de noter le langage par le biais d'idéogrammes constitués d'encoches faites par un roseau sur une feuille d'argile a produit les premiers documents : les tablettes mésopotamiennes. La conjonction d'enquêter pour un journaliste, avec la disponibilité, au XXe siècle, d'une caméra et d'une bande vidéo produit un autre type de document : le reportage audiovisuel. On constate que le support utilisé doit beaucoup aux possibilités techniques d'expression du moment et imaginer quelles autres formes ce document aurait pu avoir dans un autre contexte à une autre époque [1].

Le document est en quelque sorte une « sécrétion humaine » formatée par le niveau de développement technique ou technologique de chaque époque [2]. On peut noter une interaction entre l'homme qui cherche des moyens d'expression et les progrès techniques qui, favorisant l'expression, lui suggèrent de nouvelles idées. C'est ainsi que l'écriture, qui était d'abord transcription du langage s'est étendue à la capture du réel, via la photographie, l'enregistrement du son et de l'image, l'enregistrement d'ondes ou d'autres signaux de la nature qui sont, codés, lisibles ou interprétables par l'homme.

L'histoire des techniques d'écriture, dont le dernier avatar en date est le numérique, conditionne non seulement le support physique mais aussi la présentation de l'information, liée également aux possibilités

techniques du moment. Ainsi, une même activité humaine va générer, selon le temps auquel on se place, des documents physiquement différents, mais avec un contenu de même nature. Les exemples sont légion : le manuscrit littéraire médiéval devient livre imprimé ; le passeport du XIXe siècle comporte une description littéraire des visages (« front haut, nez plat, yeux marron... ») avant de bénéficier de la technique photographique, noir et blanc puis couleur, mais c'est toujours un passeport, c'est-à-dire un document de police pour la circulation des personnes ; le dossier médical au début du XXe siècle, lorsqu'il existe, est réduit à la narration de symptômes et du diagnostic, puis il s'enrichit d'autres documents résultant d'examens médicaux qui certes ne sont pas produits par le langage mais sont codifiées par l'homme et lisibles par lui, et c'est toujours un support écrit qui accompagne l'action de soigner un patient.

Ce qui est exprimé, le processus qui donne lieu à une trace écrite, tient à la nature de l'activité humaine et présente au fil des millénaires une stabilité remarquable autour de deux grands axes : administrer et communiquer, ou encore agir et créer. Dans le premier cas, le document accompagne et trace un acte : édicter des lois, vendre des terres, recenser des populations, juger, etc. Dans le second, l'objet écrit est sa propre finalité : manuscrit littéraire, dictionnaire, affiche, carte postale, film, etc. Mais dans les deux cas, que le processus de création soit organique ou qu'il s'agisse d'une œuvre, le document a un support, un contenu, une date et un contexte objectif de création. C'est un objet unique, inscrit dans le temps et l'espace.

La matérialité

A partir de quel moment peut-on dire qu'un document existe ? Qu'est-ce qui fait que l'on peut parler d'un document ? Qu'est ce qui précède la production d'un document ?

Avant le document, il y a l'information vive qui ne laisse pas de trace ou seulement une impression, un souvenir, qui n'est pas lisible de manière univoque comme l'est l'écriture. Ce qui fait le document, c'est d'abord sa matérialité, c'est-à-dire l'existence physique du support sur lequel est fixée l'information, matérialité qui fait que, tant que dure le support, on peut revenir à l'expression passée, à ce qui a été écrit, le lire puis l'interpréter.

Sans s'appesantir directement sur la notion de document, les philosophes l'introduisent volontiers dans leur réflexion sur la mémoire. Pour Michel Foucault, les documents sont des « traces laissées par les

hommes ». « L'énoncé, expose-t-il, est toujours donné au travers d'une épaisseur matérielle, même si, à peine apparue, elle est condamnée à s'évanouir. ». Plus loin, sans utiliser le mot « document », l'auteur de *L'Archéologie du savoir* donne une définition qui lui correspond tout à fait : « on pourra appeler formulation l'acte individuel (ou à la rigueur collectif) qui fait apparaître, sur un matériau quelconque et selon une forme déterminée, ce groupe de signes : la formulation est un événement qui, en droit au moins, est toujours repérable selon des coordonnées spatio-temporelles, qui peut toujours être rapporté à un auteur, et qui éventuellement peut constituer par elle-même un acte spécifique... » [3]

Paul Ricœur décrit le document au travers de l'archive mais c'est bien la matérialité du document qui est soulignée : « L'archive se présente ainsi comme un lieu physique qui abrite le destin de cette sorte de trace que nous avons soigneusement distinguée de la trace cérébrale et de la trace affective, à savoir la trace documentaire » [4]

Le point fondamental de la notion de document est l'intervention humaine dans la matérialisation du message exprimé, le geste d'expression et de fixation de l'information par l'auteur. Dès lors, cette information peut être lue et comprise par quelqu'un d'autre, même si elle ne l'est pas effectivement, ou qu'elle ne l'est pas tout de suite. L'existence d'un lecteur ou d'un spectateur n'est pas une condition indispensable à la définition de document, celle de l'auteur oui. Ainsi, un journal intime est objectivement un document dès l'écriture de sa première page, même si son auteur le cache, même s'il n'est jamais publié, même s'il est détruit sans avoir été lu [5].

Les notions de validation et de transmission n'entrent pas *a priori* dans la définition initiale de document ; il s'agit là d'étapes postérieures à la matérialisation de l'écrit et qui conduisent d'ailleurs souvent à une nouvelle dénomination du document : publication, ouvrage, dossier, etc.

Pour les supports d'information dans le domaine de l'administration (par opposition au domaine de la communication), la langue anglaise traduit très bien, dans le couple « document/record » cette étape de passage de la notion d'un objet d'information produit par un auteur dans son activité quotidienne, lequel peut être modifié ou complété (« a document ») à un objet d'information sanctionné par une instance supérieure à son auteur comme un tout arrêté à une date donnée et présentant une valeur de preuve, de référence ou de témoignage pour un groupe (« a record ») [6]. Certains documents ne deviennent jamais des records [7].

Le même processus chronologique existe dans le domaine de la communication ; on le voit par exemple au travers des deux mots « manuscrit » et « livre imprimé » à partir du XVI^e siècle. Le manuscrit est objectivement un document même s'il ne devient jamais livre.

Le document-trace peut être créé de toutes pièces ou à partir d'éléments existants, que ce soit d'autres documents ou des éléments naturels tel qu'un prélèvement sanguin ou une carotte glaciaire. Dans ce cas, ce qui crée le document n'est pas l'échantillon en tant que tel mais la fiche descriptive qui l'accompagne, c'est-à-dire une « méta-information » qui l'explique et le rend intelligible. De ce point de vue, dans l'exemple célèbre de l'antilope qui lorsqu'elle est se trouve enfermée dans un zoo, constitue un document [8], ce n'est pas l'antilope seule qui serait un document, mais l'antilope avec sa fiche dans le catalogue du zoo, ou mieux la fiche avec l'antilope comme pièce jointe.

Le devenir du document après sa production dépend de sa capacité intrinsèque à affronter l'usure du temps, de sa complétude informationnelle vis-à-vis de son éventuelle utilisation ultérieure, de sa sécurité extrinsèque (ni accident de conservation ni dégradation humaine).

Lorsque les documents revêtent pour leurs auteurs ou pour les institutions une valeur de mémoire légale ou culturelle, leur conservation, préventive et curative, est organisée au nom de la collectivité ; c'est la justification des centres d'archives et des bibliothèques. A noter qu'à ce stade, le simple mot « document » est remplacé par des noms plus techniques et plus précis : livres, collections, archives.

2.2 Le document-source

Une référence pour le lecteur

Dans le langage courant, le mot « document » évoque surtout l'idée d'autorité de l'information portée sur le support, de référence pour l'utilisateur, de valeur de l'information comme source de connaissance pour le lecteur. Le terme « document » a là une connotation un peu solennelle qui souligne l'étymologie du terme dans sa valeur d'enseignement (latin *docere*).

On trouve dans les catalogues d'éditeurs la rubrique « Documents » proche mais distincte de celle de « Témoignages » : les témoignages s'adressent à un public ciblé ; les documents ne sont pas destinés *a priori* à un public particulier mais ils sont proposés à tous ceux qui veulent

apprendre. La presse qualifie volontiers de « document » un ouvrage, une lettre ouverte, un rapport ou tout autre support d'information dès lors qu'il est appréhendé dans sa potentialité d'enseignement, voire d'enseignement universel.

A Pécuchet qui trouve inepte de lire Paul de Kock et de vieux ermites de la Chaussée d'Antin, Bouvard, dont le but est la connaissance, rétorque « Mais par la suite, ce sera fort curieux, comme documents » [9].

Dans cette acception du terme, la matérialité, bien que toujours présente, semble cependant reléguée au second plan au profit du contenu. De sorte que la définition : « un document est constitué d'information portée par un support » [10] peut s'entendre de deux façons : une information qui est portée sur un support par son auteur qui est aussi son premier utilisateur ; et une information qui a été portée hier sur un support et qui est disponible aujourd'hui comme source d'enseignement pour tout utilisateur autorisé. Une autre définition de document : « information enregistrée ou objet qui peut être traité comme une unité » [11] insiste sur la cohérence de l'information mais cette cohérence peut s'appliquer aussi bien à l'unité d'information produite qu'à l'unité d'information exploitée.

L'antilope du zoo dépourvue de sa fiche d'inventaire peut être un document pour celui qui sait en tirer une information sans être un document-trace au sens où on l'a vue précédemment. C'est la lecture, et non plus l'écriture, qui fait le document et, dans ce sens, un artefact, un monument visible et palpable peut être un document [12]. Dans ce cas, les techniques d'écriture ne sont pas en cause.

Le questionnement

Le document source s'apprécie relativement à celui qui le considère comme tel, c'est-à-dire comme une source d'information, et non plus comme un document dans l'absolu ou relativement au périmètre d'action de son auteur. Alors qu'un document-trace a forcément une date de création, la connaissance, tirée de l'information, est a-temporelle dans le sens où elle appartient toujours au temps de celui qui la manipule.

L'utilisateur du document n'est pas objectif. Il a une demande d'information, exprimée ou intuitive, qui lui fait regarder tel et tel support d'information comme sources pour l'acquisition de connaissances. On peut dire que pour l'utilisateur, le document est la conjonction d'une interrogation et d'une information disponible.

Paul Ricœur souligne, pour l'historien, l'importance de l'interrogation dans la relation avec les documents : « Les documents ne parlent que si on leur demande de vérifier, c'est-à-dire de rendre vraie, telle hypothèse » ; « pour l'historien, le document n'est pas simplement donné, comme l'idée de trace laissée pourrait le suggérer. Il est cherché et trouvé. Bien plus, il est circonscrit, et en ce sens constitué document, par le questionnement. Pour un historien, tout peut devenir document... » [12]

Michel Foucault, dans sa réflexion sur les « énoncés » (l'information fixée sur le support), estime que les énoncés doivent être considérés « dans l'épaisseur du cumul où ils sont pris et qu'ils ne cessent pourtant de modifier, d'inquiéter, de bouleverser et parfois de ruiner ». L'existence aujourd'hui des énoncés d'hier, leur « rémanence », ne signifie pas « qu'ils restent dans le champ de la mémoire ou qu'on peut retrouver ce qu'ils voulaient dire ; mais cela veut dire qu'ils sont conservés grâce à un certain nombre de supports et de techniques matériels..., selon certains types d'institutions..., et avec certaines modalités statutaires... » [13]. C'est donc le temps de l'utilisateur du document, avec ses préoccupations, qui prime et non la reconstitution du temps passé. Cependant, pour que l'information d'hier soit disponible aujourd'hui, il faut qu'elle ait été conservée, dans sa complétude et son intégrité, soit par le hasard, soit par une action délibérée.

2.3 Relations entre traces et sources

Il est intéressant de remarquer que les deux ensembles de documents qui sont produits (documents-traces) et des documents qui enseignent (documents-sources), ne se recouvrent pas complètement : certains objets documentaires sont réellement duals et appréhendés dans les mêmes termes en trace et en source, tandis que d'autres ne sont que trace (le document jamais utilisé) ou que source (un fossile) et, plus souvent, partiellement trace et source, avec des multiples combinaisons. Pour tenter de modéliser ces relations, on peut considérer : a) que le document-trace est soit parachevé, c'est-à-dire qu'il constitue un objet sans ambiguïté, délibéré, complet, contextualisé et authentique, soit inachevé, laissant dès lors à l'utilisateur le soin de le compléter à sa guise ; b) que la source, soit s'identifie rigoureusement à un document-trace, soit emprunte à un document-trace de manière explicite, soit encore se compose d'informations qui sont autant d'éléments, de fragments de documents désolidarisés de leur auteur initial, sans référence au document-trace préexistant. On obtient avec ce raisonnement plusieurs cas de figure théoriques dont certains se recourent. On retiendra :

- identité entre trace et source ; par exemple, la lettre d'un écrivain à son éditeur utilisée comme source d'étude littéraire, ou le plan d'une usine de fabrication de matières dangereuses exploité dans le cadre du démantèlement de ladite usine ;

- document-trace partiellement utilisé comme source, ou source se référant à telle ou telle partie, passage ou composant du document-trace (un article dans un journal ou une citation) ou bien à la totalité ou à une partie d'un document inachevé (un projet de libelle dans une liasse de papiers judiciaires du XVIIIe siècle portant l'étiquette « papiers inutiles ») ;

- source centrée sur un contenu déconnecté d'une trace initiale, sans souci d'origine, d'authenticité ou de contexte de création (collection de timbres-poste, petites phrases d'hommes politiques, images de chats).

Ces trois types de relations semblent devoir fonctionner intellectuellement quel que soit le support d'écriture ou d'information, et quels que le soient le contenu informatif et l'activité qui l'a produit.

3 L'AGE NUMERIQUE

3.1 Permanences

La question est de savoir si la rupture flagrante qu'introduit le numérique est de nature à remettre en cause la définition du document en tant que trace ou en tant que source.

Le raisonnement des pages précédentes, en privilégiant l'action et sa raison d'être, c'est-à-dire le geste qui a produit le document, comme essentiel par rapport aux outils qui ont permis sa matérialisation, conduit à minimiser l'effet révolutionnaire du numérique dans l'appréhension même de la notion de document.

L'environnement numérique a fait naître de nouvelles typologies qui, si on y regarde de près, sont centrées sur l'aspect formel du document sans s'attarder sur la nature du geste qui a produit cette trace ou la catégorie de connaissance dont les documents peuvent être la source. On énumère, comme types de documents, des fichiers, des pages web, des e-mails, des enregistrements XML, etc., qui sont autant de types physiques qui ne disent rien sur la nature du contenu informatif du support ; on distingue comme « types de contenus » le texte, les images, la vidéo, le son, qui sont plus des caractères externes qu'internes du document, sans

prendre en compte non plus la valeur de l'information ni du point de vue de l'auteur ni du point de vue de l'utilisateur.

On constate donc que le numérique met en exergue la forme du document qu'il bouleverse radicalement, mais que ces bouleversements de la forme n'ont pas d'incidence majeure sur les fondements de la démarche de production de l'information (administration ou communication) et de la démarche de questionnement vers l'information disponible. Si on reprend les définitions de document de l'environnement analogique et qu'on les applique à l'environnement numérique, on constate qu'elles restent valides. Le numérique modifie les composantes du document, pas la notion de document en tant que telle, ni l'existence de documents-traces et de documents-sources.

Les choses sont cependant moins claires, à partir de ces mutations de formes, dans les relations entre traces et sources.

La question est parfois posée de savoir si une base de données est un document. Sans développer ici une typologie des bases de données, considérons simplement les bases de données héritières des registres d'antan (cadastre, catalogue de bibliothèques, registre des élèves, etc.) : l'activité d'enregistrement des propriétés, des livres ou des élèves produisait hier un registre et produit aujourd'hui une base de données ; le fait que la base de données soit ouverte et continuellement modifiée peut laisser penser que ce n'est pas un document au même titre que le gros volume relié qui l'a précédée, lequel était cependant ouvert lui aussi jusqu'au noircissement de la dernière page. Mais à y regarder de près, on peut dire que l'environnement papier réduisait abusivement la notion de document à l'unité physique représentée par un volume relié, alors qu'il s'agissait finalement toujours du seul et même document tout au long de l'activité et que c'est la contrainte matérielle du support qui faisait considérer trente documents différents là où il n'y en avait finalement qu'un, un seul registre en trente volumes, c'est-à-dire un document unique de même il n'y a qu'une base de données pour les remplacer, de laquelle base de données on peut extraire des ensembles de données correspondant aux volumes de naguère, ou d'autres ensembles.

Les réflexions autour du document numérique qui prennent en compte la valeur de preuve ou d'œuvre du document en donnent des définitions qui sont parallèles à la définition de l'ère analogique : « 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 ». [15]

On rencontre la même opinion dans le domaine de la création : « Pour complexe que soit la notion de document numérique, elle n'est pas pour autant obscure. Elle est seulement complexe dans le sens où elle suppose une architecture verticale de langages différents. [...] Dans le monde qui s'achève, l'œuvre véritable était le livre imprimé. Quelle est-elle aujourd'hui ? Le fichier propriétaire ? en code source ? en code hexadécimal ? en code binaire ? Cette question est moins complexe qu'elle ne le paraît d'abord. L'œuvre véritable est l'œuvre éditable (et non plus seulement réitérable), c'est à dire une copie transparente dont on peut modifier librement le code source. » [16].

Le document numérique en tant que source de connaissance existe tout autant que dans l'environnement analogique, davantage si on considère les possibilités décuplées de recherche et d'accès à l'information que donnent les réseaux et les moteurs de recherche, autrement si l'on prend en compte l'évolution des modalités d'identification et de critique des sources.

La multiplicité des sources mises à disposition via Internet et les possibilités techniques des outils d'accès engagent le lecteur à naviguer, à sélectionner, à comparer, à recomposer, à construire « sa » source d'information, au risque, par un excès d'opérations successives et du fait d'une certaine opacité du web en termes de provenance et d'authenticité, de perdre la trace de l'information initiale et la traçabilité de sa propre démarche.

On remarque dans l'utilisation d'Internet que l'appellation « document » semble en perte de vitesse au profit du mot « ressources » qui évoque à la fois une plus grande potentialité d'information et le caractère indirect de l'accès aux données. Il semble que les « quatre coins du document » se soient dilués dans la masse numérique, quelquefois à tort si cette absence masque l'ensemble cohérent d'où sont tirées les données et fragilise leur interprétation [17].

3.2 Mutation de la notion de support

De la matière à la chaîne logicielle

L'expression communément utilisée de « dématérialisation » des documents ou dématérialisation des supports est impropre ou du moins inexact puisque les 0 et les 1 qui codent l'information numérique sont

fixés sur des bandes magnétiques ou des disques optiques qui ont, sans doute possible, une matérialité. La dématérialisation désigne en fait « l'allègement et la mobilité croissante des supports d'inscription, la mutation des traces et la rupture des contacts physiques autrefois liés à l'échange » [18]. C'est pourquoi, il serait plus exact de dire que le numérique virtualise le document, c'est-à-dire le fait passer du visible au possible, sans en changer la quintessence et que le phénomène révolutionnaire n'est pas la remise en cause de la notion de document mais, au sein de la notion stable de document, la dilution de la notion de support dans tout un ensemble d'éléments qui supportent et qui transmettent l'information.

Par habitude de l'environnement analogique où matière et information sont indissociables, on réduit facilement le support à la matière ou aux matières d'écriture (le papier avec l'encre) en oubliant la double fonction d'un support : d'un côté, supporter l'information par la fixation de l'écriture sur la matière ; de l'autre, transmettre l'information à l'utilisateur ou au lecteur car sans la possibilité de cette transmission, le document-source n'existe pas. Cette fonction globale est représentée en français par le mot « support » (lequel tend à en masquer une partie) alors que, pour le même concept, l'anglais utilise le mot « medium » qui insiste davantage sur l'aspect de transmission et de médiation.

La particularité du numérique est d'introduire une dissociation entre l'information et le support matériel puisque ce qu'on voit à l'écran ou que l'on imprime n'est pas *stricto sensu* le document produit initialement. La notion juridique d'original, fondée sur l'indissociabilité du contenu et du support s'en trouve ébranlée. C'est bien la notion de support, par définition, que le numérique remet en cause, avec, par ricochet, les principes et les pratiques qui reposent sur une définition analogique du support.

Pour la même raison, la pratique de placer le document dans une boîte sur un rayonnage *ad hoc* pour le conserver est inadapté : réduire le support numérique au support de stockage physique est un raccourci de l'âge pré-numérique. Le support numérique n'est plus un morceau de matière mais une chaîne matérielle et logicielle qui à la fois supporte et transmet le contenu : outils d'écriture, formats d'encodage, support de stockage, matériel de restitution. Lors de la fixation de la trace comme lors de l'identification de la source, il est indispensable de prendre en compte l'ensemble de la chaîne d'éléments qui permettent la matérialisation et la visualisation des données.

On peut donc parler de glissements technico-juridiques à l'intérieur de la notion de document mais tant que l'on peut identifier des auteurs, des actes ou des œuvres matérialisées par des techniques d'écriture, on est bien en présence de documents.

Conséquences de la redéfinition du support

Les conséquences de la stabilité du support dans l'environnement traditionnel étaient une relative confiance en l'intégrité du document (encore que les faux n'aient pas attendu l'électronique, loin de là) et une relative auto-conservation des supports en tant que morceau de matière, au moins pendant un certain temps. Pour les acteurs du documents que sont les auteurs, les éditeurs, ceux qui assument le contenu et la portée des actes, les archivistes et les bibliothécaires qui assurent la conservation, les chercheurs qui doivent identifier les sources et les documentalistes qui sont spécialisés dans cette tâche, le support papier représentait quelque chose de concret, de visible, de durable (même si la matière, venant à s'altérer, exigeait une restauration) et la difficulté de gestion de ces supports résidait essentiellement dans la logistique et la description des contenus pour une meilleure mise à disposition de l'information.

Dans l'environnement numérique, avec la nécessaire transposition de la composante « support » du document, les exigences de production, de gestion et d'exploitation du document sont différentes. La première exigence pour créer la trace ou pour identifier l'information est d'avoir affaire à un objet stable auquel on puisse se fier (ce que je vois ou ce que je lis est bien ce que je suis censé voir ou lire), qui garantisse l'existence du document dans le temps et conserve l'information pour les utilisateurs potentiels. Il faut donc fixer le document d'une manière ou d'une autre pour contrer la volatilité naturelle de l'écrit numérique. La seconde exigence, si on veut que le document perdure, étant de maintenir cet objet stabilisé au cours du temps. Seule cette démarche de « fixation » dans une forme donnée à une date donnée permet de respecter la notion de document à la fois comme trace (voilà ce que l'auteur a exprimé) et comme source (voilà l'origine de cette information).

Le numérique raccourcit considérablement la durée de vie du document non assisté. Bien sûr, dès que l'on a appuyé sur la touche « enregistrer » du clavier, on peut considérer que le document existe [19] mais sans autre mesure, il n'a aucune chance de durer, dans le sens d'être lisible, ce que peut durer un document papier, même de médiocre papier.

Le défi est d'importance puisque le groupe de recherche international sur le document d'archives électronique, InterPARES, énonce comme un préalable à ses travaux : « Conserver un document électronique est à proprement parler impossible ; seule la capacité de le reproduire peut être préservée » [20].

La réponse se situe donc dans une suite d'opérations techniques et organisationnelles auxquels le message exprimé doit se soumettre pour constituer un document viable, c'est-à-dire un document complet, lisible, interprétable et durable, tout au moins pour la durée voulue par l'auteur ou le détenteur du document.

Ces évolutions technologiques s'accompagnent forcément d'une évolution terminologique, tant la forme des choses diffère. Par exemple, le modèle OAIS (Open Archival Information System) parle d'un « objet de données » qui, doté de son « information de représentation » constitue un « objet d'information » pérennisable [21]. L'objet de données est le document tel qu'il a été produit au sein du périmètre d'activité de l'auteur et l'information de représentation, qui n'est autre que les métadonnées, permet la fixation du message initial dans un objet complet et interprétable.

Si on y réfléchit, les métadonnées ont toujours existé, sous forme de caractères internes ou externes du document [22] mais leur indissociation, au sein du document, d'avec les données auxquelles elles se rapportaient, ne requerrait pas d'en faire un concept à part.

De même, la structuration des documents et le balisage des éléments d'informations est la réponse naturelle d'une part à la contrainte que constitue la perte de la lisibilité directe de l'information enregistrée sur le support, et d'autre part aux possibilités technologiques de la recherche d'une chaîne de caractères en lien avec des outils sémantiques. On peut ainsi « prendre en compte la granularité des informations qui peut être une phrase, un paragraphe, un chapitre, un document, un hyper document, etc. pour donner à l'utilisateur des réponses précises. » [23]

3.3 En quoi le numérique fragilise mais ne remet pas en cause la notion globale de document

Le numérique, en soi, périmé les pratiques de production et d'utilisation de l'information basées sur l'indissociabilité du contenu et du support. Il perturbe le mode de fonctionnement des professionnels du support d'information (texte, image ou son), qu'ils soient spécialistes de

la maintenance du support proprement dit (chaîne matérielle et logicielle), ou spécialistes de la qualification et de la description de l'information.

Rien n'interdit aujourd'hui de penser que le numérique, qui complexifie les relations entre les documents en introduisant une chaîne de composition-décomposition-recomposition de l'information à chaque intervention de l'auteur ou du lecteur, pourra aussi gérer les contraintes engendrées par ce processus. Il apparaît que la difficulté est moins dans l'aptitude à stabiliser un objet de données et à sécuriser les systèmes d'information (car ces questions ont commencé à trouver des réponses techniques), qu'à bien circonscrire et décrire les objets d'information que l'on veut manipuler, que l'on veut créer, que l'on veut consulter.

La distinction, avec les langages de marquage des documents numériques, entre les données, la structure et la présentation du document est aujourd'hui admise. Sans doute faut-il approfondir les aspects d'identification de l'information cohérente (à différents niveaux au sein d'un même document), de description de la complétude et de la fiabilité des unités d'information, de hiérarchisation des éléments constitutifs d'un document arborescent ou interactif, voire de caractérisation de l'intention de l'auteur de déclarer insécables son document ou certaines parties de son document, et de la volonté du lecteur de reconstruire ses sources selon tel ou tel critère. On touche là aux droits de la propriété intellectuelle.

L'environnement traditionnel avait conduit à diversifier les métiers en fonction de critères de forme des documents et d'activité : les bibliothécaires pour les oeuvres, les archivistes pour les actes, les documentalistes pour la recherche d'information. L'environnement numérique change la donne et tend à redistribuer les métiers autour de deux démarches fondamentales pour l'écrit numérique : l'identification au moment de sa création d'un objet de données cohérent qui sera fixé en tant que document avec toutes ses subdivisions informatives à destination du lecteur, et la maîtrise de la maintenance technique et organisationnelle du document en tant que chaîne matérielle et logicielle.

4 CONCLUSION

La notion de document est chahutée par les bouleversements de l'âge numérique mais retrouve ses marques car il y a toujours des hommes, des actes et des oeuvres lisibles et exploitables par l'homme. En tant qu'aboutissement d'une action ou en tant que point de départ d'une connaissance, le document perdure au travers des techniques et des

technologies. Il perdure dans sa dualité de trace et de source, au-delà de l'évolution propre de ses composantes culturelles ou technologiques, et principalement du support.

Le mot « document », numérique ou pas, renvoie donc soit à un objet unique portant une information tracée à un instant T, doté de coordonnées spatio-temporelles (un auteur, une date ou une période) ; soit à une information hébergée sur un support quelconque (la notion de support reste importante dans la définition de document) considérée en tant que source, valide, en fonction de la recherche subjective d'un ou de plusieurs utilisateurs/lecteurs. On pourrait du reste imaginer deux mots distincts pour désigner cette double réalité : trace matérielle créée par l'homme à travers l'écriture (au sens large) d'une part ; groupe d'informations tangibles susceptibles d'apporter la connaissance d'autre part. Le débat pourrait s'en trouver clarifié¹.

La dissociation du support et de son contenu engendrée par le numérique, avec ses conséquences sur le plan du droit et de la conservation, ne remet pas en cause la notion de document, même si elle la fragilise ; elle redéfinit toutefois radicalement la notion de support, invitant les professionnels du document numérique à s'adapter.

Il semble que la normalisation de relations entre le numérique et le document passe d'une part par une consolidation des capacités de la technologie à conserver l'intégrité des objets d'information, et d'autre part par une identification plus formelle de la granularité de l'information, en production comme en utilisation, supportée par des métadonnées *ad hoc*.

REFERENCES

- [1] M.-A. Chabin, *Je pense donc j'archive. L'archive dans la société de l'information*, L'Harmattan, 1999, p 38
- [2] M.-A. Chabin, « L'archive numérique », dans actes du colloque *Textualité et nouvelles technologies*, Montréal, Musée d'Art contemporain, 24-26 octobre 2001, p 218
- [3] M. Foucault, *L'Archéologie du savoir*, Paris : Gallimard, 1969, pp 15, 132, 140.
- [4] P. Ricœur, *La mémoire, l'histoire, l'oubli*, Le Seuil, 2000, p 210

¹ A titre d'exemple de doublet linguistique pour désigner deux facettes d'une entité *a priori* unique, on peut se risquer à citer le mouton anglais : « sheep » pour la bête sur pied et « mutton » chez le boucher.

- [5] Roger T. Pédaque, paragraphe 128 de « Document : forme, signe et médium, les re-formulations du numérique. », 08 juillet 2003. Working paper. http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/05/11/index_fr.html
- [6] Commission européenne, MoReq, *Model requirements for the management of electronic records*, 2002 - Voir <http://www.cornwell.co.uk/moreq.html>
- [7] *The concept of records, Report from the Second Stockholm Conference on Archival Science 30-31 may 1996*, Lund, 1998
- [8] M. K. Buckland, "What is a "document"?", in *Journal of the American Society of Information Science* 48, no. 9 (Sept 1997): pp 804-809
- [9] G. Flaubert, Bouvard et Pécuchet, Paris, 1881, chapitre V
- [10] Loi québécoise concernant le cadre juridique des technologies de l'information, 2001
- [11] Norme internationale ISO 15489 sur le Records management, 2001
- [12] R. Debray, "Trace, forme ou message ? ", in "La confusion des monuments", *Cahiers de médiologie*, n° 7, coordonné par Michel Melot, Gallimard, 1999
- [13] P. Ricœur, *op. cit.*, p 226
- [14] M. Foucault, *op. cit.*, pp 162-164
- [15] G. Gallezot et alii, « Normes et standards dans le processus de traitement du document numérique en biologie moléculaire », *Solaris*, n° 6, Paris, 1999 - [internet]. Consulté en janvier 2004 : <<http://biblio-fr.info.unicaen.fr/bnum/jelec/Solaris/d06/6gallezo.html>>
- [16] J.-P. Depétris, « Ce que pourrait être un art libre », [internet]. Consulté en janvier 2004 <<http://jdepetris.free.fr/load/arlib.html>>
- [17] P. Emerson, « Electronic Records still need the 'old' disciplines – a report on MER 2001 », Consulté en février 2004 <www.mybestdocs.com/pe-mer2001v2.htm>
- [18] F.-B. Huygues, in « Pourquoi les médiologues ? », *Cahiers de médiologie*, n° 6, Gallimard, 1998, p 269
- [19] J. Derrida, *Mal d'archive*, Galilée, 1995, p 46
- [20] www.interpares.org
- [21] C. Huc, « Le Modèle de référence pour les systèmes ouverts d'archivage. L'étonnant parcours d'une norme unificatrice », numéro spécial « L'archivage », dans *Document numérique*, vol 4, n° 3-4/2000
- [22] M.-A. Chabin, *Le management de l'archive*, Hermès Sciences Publications, 2000, pp 95-96
- [23] C. Berrut, « Modélisation et Recherche d'Information Multimédia. Modèles de Documents : Documents Structurés », Consulté en février 2004 <<http://www-mrim.imag.fr/presentation/ds.php>>

