Documents, ressources, données : les avatars de l'information numérique

Sylvie Lainé-Cruzel

▶ **To cite this version:**

Sylvie Lainé-Cruzel. Documents, ressources, données : les avatars de l'information numérique. Revue I3 - Information Interaction Intelligence, 2004, 4 (1). sic_00001018

HAL Id: sic_00001018

https://archivesic.ccsd.cnrs.fr/sic_00001018

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.

Documents, ressources, données : les avatars de l'information numérique

Sylvie Lainé-Cruzel

ERSICOM - Université Jean Moulin Lyon 3
laine@univ-lyon3.fr

Résumé

En s'intéressant à la notion de document électronique, on différencie deux évolutions qui sont utiles à des titres distincts : d'une part les documents, qui ont valeur de référence, dont l'administrateur veillera à préserver l'intégrité et dont le contexte de production est fortement significatif, et d'autre part les ressources, éléments évolutifs par nature, que l'utilisateur doit pouvoir s'approprier et qui doivent donc être mises à disposition de manière à ce que la production s'efface au profit d'une logique d'exploitation et de service.

Mots-clés *: document électronique, ressource électronique, logique d'usage, logique de production, revues électroniques, modes d'enrichissement.*

Abstract

By focusing on the notion of electronic document, we differentiate two evolutions which are useful to distinct titles: on the one hand are the documents, which have value of reference. Its administrator will take care to preserve its integrity and its context of production which is strongly significant. On the other hand are resources, evolutionary elements by nature, which the user must be able to appropriate and who must thus be placed at the disposal so that the production is erased, to focus on a logic of exploitation and service.

Key-words *: electronic document, electronic resource, logic of production, electronic reviews, modes of enrichment.*

1 INTRODUCTION

Toute tentative de définition ou de clarification sur ce qu'est le document numérique relève d'un défi un tant soit peu tautologique : le document numérique sera ce que nous déciderons de désigner ainsi.

Notre approche consiste ici en effet à tenter de clarifier les principes qui pourraient nous amener à décider, au sein d'un continuum d'objets numériques de natures très diverses, lesquels peuvent porter le nom de document, et à quel titre ils pourront ou non le porter. Nous essaierons d'identifier les processus et les caractéristiques qui nous permettront de décider si un objet numérique est ou non un document.

En ceci notre manière de poser le problème est un peu différente de celle qui a été adoptée au sein du réseau RTPDoc, et qui s'exprime dans le document collectif « *Document : forme, signe et médium, les re-formulations du numérique* ». En effet, la synthèse signée Roger T. Pédauque est construite suivant le principe suivant : dans un premier temps examiner ce qu'est le document traditionnel, ses fonctions, et en quoi il constitue un objet de recherche fécond ; puis, dans un deuxième temps, analyser les conséquences de sa mutation vers le numérique, et la transformation subséquente des pratiques, et des questionnements des chercheurs.

Nous allons dans notre analyse partir des fonctions portées par le document traditionnel [1], [2], pour essayer d'identifier si – et comment – ces fonctions sont portées et renouvelées par les objets numériques, et comment elles s'y expriment.

2 LE DOCUMENT TRADITIONNEL : UNE EVIDENCE PHYSIQUE

Le document traditionnel est lié de manière permanente et unique à un support physique. Ce lien inaltérable confère au document traditionnel un certain nombre de caractéristiques.

2.1 Support

Le document étant lié de manière permanente à un support physique, la stabilité du document sera assurée par la protection et préservation du support, et garantie par la stabilité intrinsèque de ce support.

2.2 Évidence

Le document a valeur d'évidence, au sens où, en tant qu'objet matériel et physique, il se donne à observer. Les fonctions essentielles qu'il porte, c'est à dire la fonction de preuve (le document « pièce à conviction ») ou de renseignement (le document comme étant « témoignage ou représentation du monde »), sont accessibles à toute personne disposant d'une grille d'interprétation adéquate.

Puisque le document se manifeste en tant qu'objet physique, il occupe une place, et donc il peut être rangé (opération que l'on décrira habituellement en disant qu'il est intégré à une collection d'objets), et ainsi il pourra être retrouvé. Il a une forme stable qui peut être préservée, autant que peut être préservé l'objet lui-même. Et enfin, pour évoquer et démontrer son existence, il suffit de montrer l'objet.

2.3 Confiance

Le document ayant valeur d'évidence lorsqu'il est physiquement présent, son existence peut être admise lorsqu'il est physiquement absent, sur un contrat de confiance qui n'est pas plus exigeant que celui qui est nécessaire pour admettre l'existence de tout autre objet physique. C'est ainsi qu'un document peut être cité ou référencé, dès lors qu'il l'est dans une citation ou référence qui donnera la possibilité de localiser l'objet physique, et ainsi de vérifier son existence réelle.

2.4 Fiabilité

Si l'on admet l'idée classique selon laquelle le document est une « preuve à l'appui d'un fait », ou qu'il est un témoignage d'un état ou d'un fait du monde, on considère que le monde est premier, et que le document est second et rend compte du premier. La validité et la valeur du document tiennent donc à la manière dont ils rendent compte : le document est-il fiable, est-il authentifiable ? peut-on être sûr qu'il n'a en aucune manière été transformé ou falsifié ?

2.5 Genèse

Si l'on considère que le cycle de vie du document s'organise autour de 3 phases fondamentales qui seraient : sa production, sa médiation et son exploitation (*i.e.* son usage), il est intéressant d'examiner celles qui sont

déterminantes pour donner au document traditionnel son statut et son identité de document.

On peut admettre avec Buckland [3] que n'importe quel objet de l'univers peut devenir document s'il est intégré à un système d'information lorsque, à ce titre, il sera décrit de manière à devenir identifiable au sein du système. Cette vision met la médiation au cœur du processus de création d'un document. Un objet quelconque (Buckland détaille l'exemple de l'antilope dans un zoo) devient document s'il est associé à une notice, qui permet de le situer dans une collection et donc de le retrouver.

Cependant, le document, dans son acception la plus courante, a été réalisé par son auteur à des fins informatives (ce n'est pas le cas de l'antilope). Il aura souvent par ailleurs suivi un processus éditorial (ou comparable à un processus éditorial) qui l'aura légitimé et validé.

L'ensemble de ce processus (qui constitue la phase de production), associé avec la médiation, est donc déterminante pour donner au document traditionnel son identité de document. Production et médiation sont beaucoup plus significatives que l'usage pour attribuer au document son statut.

2.6 Unicité

L'unicité du support physique confère à l'objet document une cohérence qui le destine à être exploité globalement : aussi, même si le document n'est pas nécessairement destiné à être lu ou exploité dans son intégralité, il est habituellement conçu de telle manière que l'utilisateur, qui aura en mains un objet global, puisse en avoir une perception globale (en termes d'organisation et de contenu).

3 L'INFORMATION NUMERIQUE : CARACTERISTIQUES GENERALES

Dans une mémoire informatique, tout est information numérique, et tout est stocké dans des fichiers : programmes, données, textes. Est aussi information numérique tout résultat fourni par un traitement ou un programme. Mais toute information numérique n'est pas document numérique, de même que tout objet n'est pas document, ni même simplement porteur d'une fonction documentaire. Sans doute

retrouverons-nous la mise en collection et le rangement au cœur du processus qui donnera naissance au document numérique, mais cette médiation elle-même adopte des formes renouvelées. (Le terme médiation est utilisé ici pour évoquer l'idée de rangement d'un objet en vue de permettre à d'autres de le retrouver plus tard : à la fois « incorporer dans une collection » et « rendre accessible »).

3.1 Production

Comme pour les objets documentaires non numériques, que nous appellerons dorénavant objets traditionnels (faute d'un meilleur terme), il semble que la plupart du temps l'auteur ait participé de manière consciente et active à la création d'un objet d'information. Nous nous intéresserons en priorité à ces objets qui ont été produits avec une intention informative.

3.2 Support

La notion de support ne disparaît pas avec l'objet numérique, mais elle se transforme. Il peut dorénavant y avoir co-existence de deux types de support : le support de stockage (mémoire), et les supports de visualisation ou de restitution (écrans, imprimantes etc). Si nous admettons qu'il n'y a pas d'objet sans support, y compris lorsqu'il s'agit d'objet numérique, nous devrons dorénavant différencier l'un et l'autre type de support. Distinction très significative lorsque la différence entre l'objet restitué et l'objet mémorisé devient de plus en plus grande.

3.3 Stabilité

Le lien avec le support ne garantit pas la stabilité de l'objet numérique.

- Le support de visualisation qu'est l'écran est fugace et volatile.
- Le support de restitution qu'est l'imprimante ne donne pas toujours lieu à un objet susceptible d'être rangé, et s'il l'est, le « document imprimé » est un nouveau document : un document traditionnel, distinct du document numérique auquel nous nous intéressions.
- C'est donc uniquement la stabilité et l'intégrité de son stockage qui garantira la stabilité et l'intégrité du document numérique.
- D'un point de vue physique, le support de stockage n'est stable que s'il s'agit d'une mémoire non modifiable (mémoire de type

ROM). Lorsqu'il s'agit d'une mémoire vive, elle peut être mise à jour à tout moment. *La stabilité, si elle constitue une caractéristique significative d'un document, ne pourra être maintenue que par un processus de contrôle volontaire et conscient.*

3.4 Valeur d'évidence

Si l'objet numérique est toujours associé à un support (son support de stockage), il ne s'agit plus d'un support physique directement observable. L'objet ne se manifeste qu'à travers des dispositifs de restitution complexes et dont la disponibilité et la fiabilité ne sont ni certains ni garantis. Le contrat de confiance qu'impliquent la citation ou le référencement s'en trouve sans doute compliqué.

4 DIVERSITE DES INFORMATIONS NUMERIQUES

A titre d'exemples, énumérons quelques informations numériques de nature diverses :

[a] une table de logarithmes archivée en format PdF

[b] un programme capable de calculer le logarithme d'un nombre X

[c] le résultat fourni par ce programme pour le nombre Xi

Si *[a]* est incorporée à une collection (et donc accessible) elle peut être citée. La stabilité du contenu du document est garantie par le format adopté. Le processus de production pourra être décrit en terme de responsabilité autoriale (source). La fonction de la table est bien de fournir « une preuve ou un renseignement ». Il nous paraît ici légitime de parler de *document numériqu**e***, d'autant plus facilement que l'objet numérique est très proche de la table de logarithme imprimée, plus traditionnelle.

[b] n'est pas destiné à être interprété par un être humain, mais à être utilisé. Il est destiné non à témoigner ou prouver, mais à rendre un service. Les informations qu'il est susceptible de mettre à disposition d'un utilisateur ne sont que potentiellement disponibles, et n'ont pas d'existence vérifiable hors d'une situation d'utilisation qui implique des conditions de réalisation. La fonction essentielle du programme est d'être utile : sa conception est entièrement tournée vers l'usage. Si nous admettons qu'un objet devient document par le processus qui lui donne

naissance (production/médiation), le programme (en particulier le programme exécutable) est un objet d'une toute autre nature, et dont la nature informative ne se manifeste que dans et par l'usage : nous proposons de l'appeler ***ressource.***

[c] n'est lié à aucun support, si ce n'est le support fugace utilisé pour une visualisation, ou le support d'impression qui fera éventuellement naître un nouveau document papier. Il n'est pas lié à un support de stockage, mais peut être reconstruit fugitivement à chaque utilisation du programme. Ce n'est pas un objet au sens où il n'a pas de matérialité propre. Or, s'il n'y a pas matérialité, il n'y a plus d'évidence physique : l'existence de l'objet ne peut plus être postulée sur le même contrat de confiance que celui dont nous parlions en 1.2. La citation ou le référencement ne sont donc plus possibles.

[d] le site Web d'une université

[e] une encyclopédie sur DVD

Les deux objets que nous évoquons ici ont des points communs : ils organisent des informations de natures documentaires classiques, se manifestant sous la forme de textes et d'images. L'un et l'autre sont alimentés par des processus contrôlés, sous la responsabilité institutionnelle de l'université pour le premier, et sous une responsabilité éditoriale pour le second. Cependant les attentes et les caractéristiques « idéales » qui s'expriment par rapport à l'un et par rapport à l'autre sont bien différentes.

Le site de l'université *[d]* se doit d'être à jour : renseigner d'une manière efficace, présenter des formations et des équipes existantes, des annuaires exploitables, etc. Sa fonction est d'être utile et de rendre des services. Sa nature est d'être *évolutive*, pour restituer une image aussi fidèle que possible d'un univers en transformation permanente. La qualité du site sera liée à sa capacité à évoluer en même temps que l'univers sur lequel il informe. Ce site est conçu pour une logique de l'usage. Nous proposons de le considérer non comme un document, mais comme *une ressource.*

Le site de notre exemple s'inscrit dans une double logique de médiation (rassemblement et mise à disposition d'informations que l'on souhaite rendre accessibles) et d'usage (il est destiné à rendre un service). La logique de production est passée au second plan : les textes et images ne sont habituellement ni signés (productions souvent anonymes) ni datés (le site « idéal » ne contiendrait que des informations parfaitement à jour, éventuellement confirmées le matin même).

Ainsi, ce site ne peut être considéré comme un document, au sens où il ne constitue pas un objet stable, et donc à ce titre il ne peut faire preuve (au sens « pièce à conviction »). Le citer ou le référencer n'aurait guère de sens, car l'objet auquel accèderait un autre utilisateur à un autre moment ne serait plus le même – seule persiste sa fonction informative, et sa valeur d'usage. Les ressources sont partageables mais elles ne le sont pas au même titre que le sont les documents : nous ne parlerons ni de citation ni de référence, mais nous pouvons dire qu'une ressource peut être évoquée, conseillée, indiquée, comme étant capable de rendre un service.

L'encyclopédie numérique *[e]* a une autre mission d'une nature plus patrimoniale : celle de collecter des productions (signées, datées, situées…). Ces productions doivent présenter des garanties d'authenticité et de non falsification. Les deux logiques dominantes à sa constitution sont les logiques de production et de médiation. En ceci, l'encyclopédie numérique a des caractéristiques comparables à celles qui sont mises en œuvre dans les documents traditionnels, et pourra être considérée comme un *document numérique*.

5 L'INFORMATION NUMERIQUE ET SES AVATARS

5.1 Terminologie

Nous proposons donc de différencier 3 niveaux d'informations numériques :

- Les données : des informations stockées sous forme numérique, considérées indépendamment de tout contexte de production et d'interprétation.
- Les ressources : nous désignerons ainsi des informations construites dans une logique de médiation et d'usage (réception), évolutives (susceptibles d'être mises à jour), et éventuellement adaptables (personnalisables). Leur fonction est d'être utiles et de rendre des services. Elles fournissent du renseignement (instantané) mais non de la preuve (au sens d'un témoignage historique ayant été préservé).
- Les documents : ils sont élaborés par des logiques de production qui leur donnent naissance, et des logiques de médiation qui les rendent accessibles. Ils ont valeur de preuve, et doivent être authentifiés et stabilisés. Les concepteurs de systèmes de gestion

de documents numériques auront pour préoccupations : l'authentification (exemples : tatouage, signature électronique), la reconnaissance d'une valeur légale des documents numériques, éventuellement la protection contre la copie (protection intellectuelle et droit d'auteur), la pérennité. Ils privilégieront les supports non modifiables et les formats « fac-similé ». Leurs priorités techniques sont celles qui sont prises en compte dans les systèmes gérant des *archives électroniques* [4].

Nous proposons donc de différencier documents et ressources en les désignant par des termes distincts ; alors même que la plupart des auteurs utilisent couramment les termes de « document dynamique » ou de « document évolutif » pour désigner des objets que nous nommerons ressources. Ce choix terminologique est justifié par les raisons suivantes :

- le terme de « ressources » est déjà largement utilisé par certaines communautés qui fondent leur travail de conception sur la question de l'usage et de la réception : on parle ainsi couramment de « ressources pédagogiques »,
- les concepteurs de systèmes d'informations, qui assument la fonction de mise en collection et médiation des documents numériques, auront à assumer selon la nature des informations contenues dans leur système des exigences et des logiques de nature si différente qu'elles en deviennent contradictoires et incompatibles. Utiliser le terme de « document » dans tous les cas nous semble donc un choix malheureux qui ne peut qu'obscurcir la discussion.

5.2 XML : vers un nouveau point de vue sur le document ?

Les évolutions les plus significatives de l'information numérique en cette dernière décennie sont apparemment les mutations liées à XML, et les perspectives qu'ouvre ce nouveau standard, du point de vue de la diversité des traitements et des recompositions d'informations qu'il permet d'envisager selon les contextes d'applications. Ainsi, avec la possibilité de décrire la grammaire d'une classe de documents (sous forme de DTD ou de schémas) apparaît la possibilité de diversifier les feuilles de style associées, et même les exploitations.

Le document numérique [5 ; 6] est ainsi, de plus en plus, perçu comme un objet nécessairement dynamique et évolutif – parfois même, on considérera que la stabilité du document encore souhaitée par certains

dans quelques circonstances n'est qu'un héritage passéiste, sans intérêt pour l'homme moderne quand il saura vivre avec son temps. Cette vision résiste-t-elle à l'examen ?

6 DOCUMENTS ET RESSOURCES : COEXISTENCE ET COMPLEMENTARITE

6.1 Le cas des revues électroniques

Interrogés sur la manière dont ils envisagent l'évolution des revues électroniques, des étudiants de maîtrise en documentation de l'Université Lyon 3 répondent spontanément et avec une belle unanimité que le principe de rassemblement d'articles dans des publications périodiques en volumes numérotés disparaîtra bientôt, car ce rassemblement en numéros sera rendu caduc par les nouvelles possibilités d'accès directs et thématiques aux articles, en fonction de leur contenu ou grâce aux méta-données qui les décrivent – le rassemblement des articles deviendrait alors un processus purement dynamique résultant d'une requête, dans une logique de l'usage.

Et pourtant, les revues électroniques [7 ; 8] continuent, sans nécessité technique ni physique, à publier (même lorsqu'il n'y a plus la moindre impression sur papier en parallèle) des volumes virtuels, et à le faire selon une périodicité régulière. Simple perpétuation de principes qui n'ont plus de vraies raisons d'être, ou faut-il chercher une autre explication ?

Notre hypothèse relative aux revues électroniques est qu'elles ont la même fonction que les revues traditionnelles, c'est-à-dire un rôle essentiel de validation et d'archivage du travail scientifique. La fonction de production est ici extrêmement importante (elle se manifeste par le fait qu'apparaissent toujours de façon claire l'identification des auteurs, de leur appartenance, la date de validation et parfois même de rédaction des articles, et de manière générale, de multiples facteurs décrivant les conditions de production des articles de la revue). Les articles eux-même sont généralement présentés dans un format interdisant toute intervention sur le contenu (modification, extraction de fragments etc.).

Nous avions émis des hypothèses relatives au document traditionnel : que sa nature lui permettait d'être aisément cité ou référencé (cf 1.3), que sa stabilité – et sa pérennité – pouvait être garantie (cf 1.4), et qu'en tant

qu'objet complexe il devait maintenir une cohérence de construction (cf 1.7).

S'il est facile dans un système d'information de protéger un article en l'archivant dans un format qui ne permette que de restituer des fac-similés, il est plus difficile de protéger globalement une collection qui exprime une *politique éditoriale*. Conserver les principes classiques de la revue permet de garantir et de protéger la collection et sa politique de publication, en stabilisant la trace du processus d'incorporation des documents dans la revue, au delà de la protection des documents eux-mêmes. Lorsque le sommaire du numéro XX publié dans le mois MM de l'année AA aura été établi, il deviendra impossible d'y intégrer a posteriori une production, tout comme il deviendra impossible d'en supprimer une. Le contrat de confiance est consolidé, et le processus de citation et de référencement peut s'établir en toute sérénité.

Sans vouloir évoquer ici la question de la granularité de l'information, qu'elle soit ressource ou document, nous pouvons à coup sûr évoquer le fait que se créent de nombreux objets complexes (collections, ensembles, groupes) dont la logique doit, elle aussi, être prise en compte.

Ainsi, l'objet complexe que nous venons d'évoquer en parlant de revues électroniques semble être un objet de type document à la fois au niveau élémentaire (chacun des articles publié dans la revue est un document) et au niveau des objets composés qui les rassemblent (chaque numéro de la revue est un document). Le concepteur du système adopte ici une logique propre à l'objet document : garantir sa non-falsification et sa pérennité, et garantir sa stabilité (l'ajout de nouveaux numéros est possible, mais la suppression de numéros déjà parus ou leur modification est impossible).

6.2 Bases de données documentaires

Le gestionnaire d'un système d'informations documentaires ou le gestionnaire d'une bibliothèque virtuelle pourront au sein d'un même système gérer des documents (stables) et ressources (évolutives), mais l'ensemble – i.e. le système d'information - sera conçu dans une logique de service : la logique d'usage y sera plus significative que la logique de production. Pour mieux s'adapter aux attentes des utilisateurs, les concepteurs du système pourront être ainsi amenés à fragmenter, articuler et réorganiser les contenus, donc à transformer parfois les documents en ressources pour permettre aux utilisateur d'accéder à des renseignements ou des connaissances.

Il semble que la nécessité de maintenir l'existence d'objets ayant une identité stable, qui aient valeur de preuve et puissent être cités, ne soit pas remise en cause pour autant. Auquel cas il y aurait nécessaire co-existence entre deux formes différentes d'informations :

- la forme initiale, liée à un contexte de production, figée et authentifiée, conservée dans un format de type fac-similé, gardant et protégeant le statut du document,
- la forme dérivée, adoptant un statut de ressource, structurée dans un format tel qu'XML, apte à se manifester dans des formes diverses et adaptées à l'usage, et donc liées à des contextes de réception.

Ainsi donc un système d'information fonctionnant dans une logique de ressources peut incorporer des documents et des ressources.

Le projet Profildoc développé à l'Université Lyon 1 [9 ; 10] en est un exemple : le système de recherche d'informations y est conçu de telle manière que les documents intégrés puissent être fractionnés et filtrés de manière variable en fonction du profil de l'utilisateur et de son activité, ce qui amènera à privilégier certains documents ou parties de documents comme étant plus exploitables que d'autres selon les circonstances. Les résultats présentés sont des recompositions et réorganisations de fragments de documents entre lesquels la navigation est personnalisée par la présentation privilégiée de certains liens [11].

Il s'agit donc selon notre vocabulaire d'un système fonctionnant dans une logique de ressources. Mais dans ce système, la présentation personnalisée des fragments significatifs doit toujours permettre de retourner à un document fac-similé, pour laquelle la logique de conception et d'exposition choisie par l'auteur et le respect de la logique de production sont garantis.

6.3 Systèmes d'informations et preuve

De même que les documents peuvent s'intégrer dans des systèmes de ressources et parfois devenir eux-mêmes ressources, il est probable qu'une autre possibilité se fait jour : celle qui consiste à produire des documents à partir d'un système de ressources.

Le principe de l'opération consiste ici à photographier (et donc figer) l'état du système ou d'une partie du système à un moment donné, d'en authentifier les conditions de production et de faire légitimer le document résultant par un expert habilité à le faire.

Cette démarche peut avoir un intérêt dans de multiples circonstances :

- celles où un document est produit dans un cadre collaboratif complexe (cf *groupware* et *workflow*) qui aboutit à un moment donné à une production stable ayant valeur de référence,
- celles où des informations diverses alimentent en continu le système, mais où certaines étapes décisionnelles obligent à valider l'ensemble des informations connues en les enregistrant pour leur donner une valeur officielle, ou conforme à ce qu'exigent certaines démarches qualité.

Ainsi, par exemple, le combustible nucléaire est soumis à un contrôle extrêmement important imposé par le législateur : chaque gramme de matière fissile est suivi de sa production jusqu'à son retraitement. La traçabilité du combustible nucléaire traité dans les centrales d'Electricité de France est ainsi prise en charge par l'alimentation de différents systèmes d'informations.

Mais toute opération de chargement ou déchargement ne sera autorisée que lorsque des documents retraçant l'historique de la vie du combustible auront été produits à partir des systèmes d'informations, et que ces documents auront été validés (et éventuellement annotés) par des experts.

7 HYPOTHESES SUR LES MODES D'ENRICHISSEMENT

Les différences de principe que nous avons soulignées entre ressources et documents mettent en valeur des différences essentielles dans la construction et l'éventuel enrichissement de l'un et de l'autre.

Le document a une structure que l'auteur peut éventuellement revendiquer comme étant significative. Cette structure peut par exemple avoir une fonction argumentative et rhétorique qui permet à l'auteur d'exprimer des jugements, idées ou prises de position [12]. Le document peut ainsi revendiquer une subjectivité qui se justifie par le fait que le discours produit est situé[13]. Les idées qui y seront développées pourront être protégées au titre de la propriété intellectuelle et de la protection du droit d'auteur.

Un objet-document complexe ne peut s'enrichir que par des ajouts raisonnés et contrôlés, c'est à dire soumis à une politique de production autoriale ou éditoriale.

La ressource revendique l'objectivité – elle doit « dire le vrai », au sens positiviste. Cet objectif de vérité conduit à un politique d'évolution qui peut se faire par agrégation (empilement de nouvelles informations), modifications (corrections et mises à jour) et éventuellement suppressions (des informations périmées). Dans un tel système les ressources sont fractionnables, et la structure peut être construite à la réception. Leur protection relève essentiellement du droit d'usage (*copyright*) .

8 CONCLUSION

Il semble donc que l'on puisse identifier deux formes différentes de « documents numériques », l'une que nous baptisons document, et l'autre que nous baptisons ressources.

Ces deux formes sont destinées à co-exister, car elles semblent toutes les deux nécessaires. Mais elles le sont à des titres différents. Aucune des deux ne se réduit à l'autre, cependant chacune des deux peut – sous certaines conditions – être engendrée à partir de l'autre.

Nous souhaitons cependant insister sur le caractère fondamentalement différent de ces deux types d'information, au sens où le médiateur (qui les mettra en système et en collection) aura à prendre en compte des logiques, des attentes et des critères de qualité contradictoires et incompatibles, selon qu'il sera dans une logique de ressources (logique d'usage) ou de document (logique d'archivage).

- S'il est dans la logique du document, il aura à préserver et à garantir la traçabilité d'une production et la stabilité de son résultat, et maintenir une cohérence définie à la production.
- S'il est dans la logique de la ressource, il devra garantir l'évolution, faciliter l'appropriation, et permettre que la cohérence soit définie par l'usage et l'utilisateur (reconstruction).

Enfin, si la définition du document traditionnel disait qu'il avait « valeur de preuve ou de renseignement », il semblerait que la grande mutation du numérique aboutisse de fait à deux formes distinctes : la ressource, qui aurait valeur de renseignement, et le document, qui aurait valeur de preuve. Le nouveau défi pour les médiateurs, qu'ils soient administrateurs, rédacteurs, documentalistes ou gestionnaires, sera probablement dans l'articulation entre l'un et l'autre.

REFERENCES

[1] H. Fondin. *Le Traitement numérique des documents*. Paris : Hermès, 1998.

[2] S. Fayet-Scribe. Chronologie des supports, des dispositifs et des outils de repérage de l'information. *Solaris* 1997, Dossier n°4 : « Le savoir et ses outils d'accès : repères historiques », [en ligne], 1997 http://bibliofr.info.unicaen.fr/bnum/jelec/Solaris/d04/4fayet_0intro.html

[3] M. K. Buckland. What is a « document » ? *Journal of the American Society for Information Science*, vol 48(9), pages 804-809, 1997

[4] T. Thomas. Archives in a New Paradigm of Scientific Publishing. *D-Lib Magazine* May 1998**,** Physical Review Online Archives (PROLA), 1998 <http://www.dlib.org/dlib/may98/05thomas.html>

[5] J.-L. Duchet. La nécessité d'une norme commune pour la recherche sur corpus : la Text Encoding initiative (TEI) et le langage de balisage SGML. *Textes, documents et nouveaux médias.* Poitiers : Atlantique, pages 54-62, 1998

[6] H. Hudrisier. SGML, HTML, XML : l'ère des machines grammatologiques. *Passerelles*, Numéro spécial Recherche Paris 8, vol 24, pages 42-44, 1999

[7] A. Mahé. *Les usages des revues scientifiques en ligne par les chercheurs en sciences exactes*. Thèse de doctorat, Université C. Bernard Lyon 1, 2002.
<http://www.enssib.fr/bibliotheque/documents/theses/mahe/mahe.pdf>

[8] L. Halliday et C. Oppenheim. Developments in digital journals. *Journal of Documentation,* vol 57, pages 260-283, 2001.

[9] S. Lainé-Cruzel. Profildoc : filtrer une information exploitable. *Bulletin des Bibliothèques de France* – vol 5, pages 60-65, 1999

[10] S. Lainé-Cruzel, T. Lafouge, J.P Lardy et N. Ben Abdallah. Improving information retrieval by combining user profile and document unit. *Information Processing and Management*, vol 32, n° 3, pages 305-315, 1996

[11] M. Ben Romdhane. *Navigation dans un espace textuel : accès à l'information scientifique*. Thèse de doctorat, Université Lyon 3, 2001.

[12] M. Foucault. *L'archéologie du savoir*. Paris, NRF Gallimard, 1969.

[13] K. Devlin. *Logic and information*. Cambridge University Press, 1991.