

INNOVATION SCIENTIFIQUE ET TECHNOLOGIQUE

Le document numérique, cœur de l'innovation dans les TIC ?



Stéphane Bresson

Au fil des siècles, le document a joué un rôle fondamental dans le développement des civilisations. Les innovations majeures qui rythment son histoire ont entraîné des bouleversements fondamentaux dans la diffusion des savoirs et dans la structure des organisations. L'essor de la société de l'information et de la connaissance entraîne le document dans l'ère du numérique ; mais avec cette nouvelle évolution, devient-il pour autant la clé de voûte de l'innovation dans notre société ? En Basse-Normandie, un regroupement d'entrepreneurs, de chercheurs, d'enseignants et de technologues tentent d'apporter quelques éléments de réponse.

Stéphane Bresson, docteur en physique et ingénieur en instrumentation, est chargé de mission au Centre des Technologies Nouvelles de Caen

L'invention de l'écriture, il y a 5 000 ans, délimite la transition entre Préhistoire et Histoire. Vers 1450, une seconde révolution technologique marque le passage du Moyen-Âge à la Renaissance : l'imprimerie, jouant un rôle considérable dans la baisse du coût de production du livre en Europe et élargissant très significativement sa diffusion. Aujourd'hui, l'information et les échanges se dématérialisent avec l'irruption des technologies numériques. Le document change de nature, impliquant d'importantes retombées économiques, sociales, juridiques, culturelles, voire politiques : circulation des informations facilitée, réduction du coût marginal dans une économie de l'information abondante et globalisée ou encore apparition de nouveaux métiers.

UNE DÉFINITION COMPOSITE

Un travail remarquable a été mené par un collectif de chercheurs, le « RTP Doc »[1], pour donner une définition au document numérique selon trois dimensions (voir tableau). Le document est d'abord un objet manipulable, une **forme**, composé d'un support et d'inscriptions. Qu'il soit directement créé sous forme électronique ou numérisé à partir d'un objet physique, le document numérique se définit par sa construction (structure, données, mise en forme) et par ses modes d'accès (localisation, temporalité ou matériels de restitution). Il est aussi considéré comme un ensemble de **signes**, le sens et les connaissances associées s'ajoutant aux inscriptions. Les principales illustrations sont le web 2.0 et le web sémantique. Il est enfin vu comme un **médium** : on adjoint à la forme et au sens une légitimité sociale, économique, légale...

Définition schématique du document

	Document physique	Document numérique
Objet (« forme »)	Support + inscriptions	Données + structure
Sens (« signes »)	Inscriptions + sens	Texte informé + connaissances
Vecteur (« Médium »)	Inscriptions + légitimité	<ul style="list-style-type: none"> • Texte + procédure • Publication + accès

Pour compléter cette définition, il est important de distinguer la communication organisationnelle de la communication médiatique. La première est associée à la gestion ou au management et est liée à des procédures ; dans la seconde, le document s'évalue par sa notoriété ou sa valeur patrimoniale. Il faut également considérer la structure des documents. Si une partie d'entre eux sont essentiellement composés d'informations structurées, voire de données (facture = référence + objet + expéditeur + destinataire + somme + date + ...), près de 80 % ne le sont pas a priori (plaquettes commerciales, mails, films...). Toutefois, ils peuvent sous-tendre une organisation sous forme de métadonnées à l'image d'un livre composé d'un titre, de chapitres, d'un sommaire constitués de paragraphes, de phrases, de mots... Enfin, constatons que le document numérique est, par nature, multimédia. En effet, au-delà du texte, il contient aussi des images ou des sons, le tout de manière fixe ou animée ; songeons par exemple à un film avec des sous-titres.

PASSER DU PAPIER AU NUMÉRIQUE ?

Force est de constater que le document est amené à passer fréquemment du monde électronique au monde atomique : un texte conçu de manière numérique peut être imprimé pour être transmis à un autre utilisateur qui le numérise à nouveau pour le traiter... et ainsi de suite ! Il est d'ailleurs très probable que

la fonction la plus utilisée d'un éditeur de texte reste la fonction « imprimer » ! La numérisation du document engendre une forte activité d'innovation, ce qui nous conduit à poser les deux questions suivantes : la dématérialisation permet-elle de reproduire tous les actes qui semblent naturels avec un document physique ? Engendret-elle de nouveaux usages, de nouvelles valeurs au document ?

Plus concrètement, voici quelques interrogations qui illustrent mutation et dualité du document. Est-il possible de traiter à l'identique un manuscrit et un bordereau, de dématérialiser toutes les procédures de gestion de l'entreprise, par exemple un circuit d'approbation du courrier ? Une machine peut-elle, comme l'esprit humain, retrouver facilement un texte dans une image, identifier un comportement dans une scène animée ou réaliser une lecture

Un cluster Document Numérique en Basse-Normandie

La Basse-Normandie dispose dans le domaine des TIC de compétences et savoir-faire parfois méconnus : projets de recherche technologique, entreprises innovantes, formations dédiées, recherche publique ou dispositifs d'accompagnement. Si la labellisation du pôle de compétitivité TES (transactions électroniques sécurisées) a mis en valeur un premier domaine d'excellence, il en existe d'autres. Depuis un an, le Centre des Technologies Nouvelles a lancé, avec plusieurs partenaires, la construction d'un cluster dans le domaine du document numérique. Un premier état des lieux sera disponible fin 2007 grâce à une étude conduite par le conseil économique et social régional.

Quelques exemples de solutions innovantes développées en Basse-Normandie

- Les chercheurs caennais ont permis de retrouver et d'associer des informations textuelles avec des images ou des cartes conduisant à une indexation d'informations géographiques localisées dans les territoires.
- Pour l'analyse des manuscrits de Flaubert, leurs collègues haut-normands du LITIS ont développé des outils d'analyse des biffures, ratures et autres hachures, permettant aussi le traitement automatique de bordereaux de remise de chèques.
- Dans le cadre d'un projet Eureka, la société ADCIS a développé un système de vidéosurveillance avec des applications de détection d'intrusion, de comptage et de suivi de personnes dans un réseau de caméras pour assurer la sécurité des personnes et des biens, contrôler des trafics ou surveiller des lieux publics.
- Grâce au *cahier numérique*[®] développé par les papeteries Hamelin (marque Oxford) et à un stylo numérique, il est possible d'écrire manuellement sur du papier et de disposer immédiatement de cet écrit dans un traitement de texte, un mail, un « post-it » électronique ou encore un tableur.
- Les GED proposées par des sociétés comme Scanfile, Asterion ou Eurodoc gèrent automatiquement la circulation d'un flux de courrier et son circuit d'approbation.
- À travers une plateforme de services, la société Gage aide les organisations en réseau à mettre en œuvre de nouveaux services hyper-personnalisés pour l'approvisionnement, la distribution et la communication auprès des points de vente et des clients finaux.

en diagonale ? Un document numérique signé électroniquement aura-t-il la même valeur morale qu'un document papier signé manuellement ? Pouvons-nous écrire électroniquement aussi aisément qu'à la main ? Saurons-nous à l'avenir déambuler dans une bibliothèque numérique comme nous le faisons dans notre bibliothèque de quartier ?

CARACTÉRISATION DU DOCUMENT NUMÉRIQUE

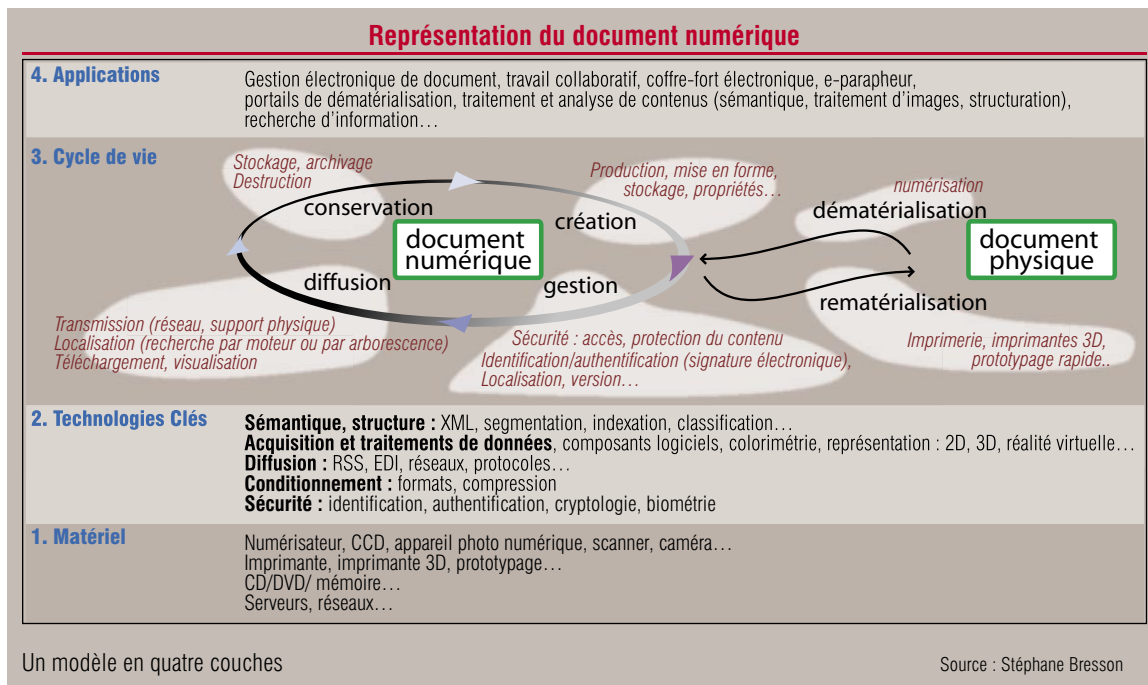
Répondre à ces questions nécessite une représentation du document numérique. Pour ce faire, nous proposons un modèle fondé sur une approche similaire au modèle OSI*, restreint ici à quatre couches (voir encadré). Paradoxalement, la première couche est celle des **matériels**. En effet, l'utilisation d'un document immatériel repose en grande partie sur des moyens physiques tant pour sa numérisation que pour son support, son transport, son affichage ou sa rematérialisation.

La deuxième est composée des **technologies clés**. Objet de nombreux travaux pluridisciplinaires de R&D dans des domaines comme la structuration, l'indexation, la sémantique, la représentation multidimensionnelle, le transport, le conditionnement, la sécurisation, ces technologies sont nécessaires au parcours du document sur son cycle de vie (couche 3) et à l'élaboration d'applications (couche 4).

La troisième couche, **le cycle de vie**, est composée de quatre étapes distinctes : la création, la gestion, la diffusion et la conservation [2]. La **création** d'un document a deux origines : soit un opérateur disposant d'outils logiciels, soit une chaîne de capture totalement numérique. Cette étape inclut aussi la mise en forme, le stockage intermédiaire et la détermination des propriétés du document, par exemple l'auteur, les mots-clés ou la date. Dans la phase de **gestion**, le document va ensuite acquérir ses caractéristiques fonctionnelles : niveaux de sécurité, authentification, localisation dans le système d'information, numéro de version... Vient ensuite sa **diffusion** avec deux modes distincts. En « *push* », il est envoyé volontairement à des destinataires par un système de transmission. En mode « *pull* », il est téléchargé et visualisé par l'utilisateur après avoir été localisé. Enfin, il doit être **conservé**, c'est-à-dire sauvegardé pour pouvoir être modifié, archivé en cas de stockage définitif à des fins de mémorisation ou de preuve, ou supprimé.

La quatrième couche recouvre **les applications** pour lesquelles nous distinguerons deux catégories. La première regroupe les outils de traitement du contenu. La seconde englobe la gestion et représentation de masses de documents avec la gestion électronique de document (GED) qui gère les flux, le stockage ou

OSI : le modèle d'interconnexion en réseau des systèmes ouverts de l'Organisation internationale de normalisation fondé sur la norme ISO 7498 est globalement intitulée « Modèle basique de référence pour l'interconnexion des systèmes ouverts (OSI) ».



encore le classement, sans oublier les portails d'information, la gestion des connaissances, les bibliothèques numériques et d'une manière globale, le web 2.0.

DES USAGES MULTIPLES

La chaîne graphique, et plus particulièrement l'édition, est, par nature, la première filière touchée par la numérisation. Si l'introduction du numérique rend possible la publication d'ouvrages à l'unité, elle nécessite aussi une réflexion sur la multimodalité. L'apparition du livre électronique est un bon exemple de rupture dans les métiers traditionnels dans le secteur. Le domaine de l'administration, du social ou de la citoyenneté est probablement le plus dynamique actuellement avec le déploiement de la e-administration : téléprocédures, télédéclarations, téléservices ou, dans la santé, le dossier médical personnel (DMP).

Du côté des entreprises, la dématérialisation des processus est aussi bien engagée. Après la bureautique, elle touche maintenant la gestion financière et les ressources humaines avec de nombreuses initiatives autour de l'électronisation des documents qui, en raison du temps légal de conservation, génèrent des grands volumes d'archives papier : factures, fiches de salaires... Pensons aussi au marketing avec des applications prometteuses dans la gestion de la relation client. Les échanges inter-entreprises sont également concernés avec les échanges de données informatisés (EDI), ou plus récemment l'« ebXML ».

Pour le secteur de la culture ou du patrimoine, la numérisation facilite la disponibilité des œuvres (livre, films, journaux...) à tout instant et à distance, éventuellement dans des bibliothèques numériques, et démocratise ainsi l'accès au savoir. D'autres applications majeures peuvent encore être identifiées, comme le droit et le juridique avec la « notarisation » des documents.

Qu'il s'agisse du papyrus utilisé par l'antique scribe égyptien, du parchemin du moine copiste ou du billet à ordre au XIX^e, les évolutions techniques ont donné au document une place centrale dans la société : mémoire, preuve ou titre de transaction. Aujourd'hui, sa numérisation amplifie encore ce phénomène.

Innover dans ce domaine implique donc de considérer tout à la fois les enjeux techniques, opérationnels, humains, économiques, voire sociétaux avec une approche d'éco-développement. Il s'agit, entre autres, de considérer l'interopérabilité, la pérennité, la traçabilité et la sécurité des solutions, de tenir compte d'un éventuel dédoublement des traitements en raison de la dualité papier / électronique, de s'attacher à l'appropriation par l'utilisateur, de se préoccuper de la formation ou de la résistance au changement, de bien estimer les réductions des coûts et le retour sur investissement, d'approfondir la gestion des connaissances ou de l'immatériel.

Centres de compétences :

- * Réseau thématique pluridisciplinaire DOC, départements STIC et SHS du CNRS, <http://rtp-doc.enssib.fr>
- * APROGED, Association des professionnels du document numérique, www.aproged.org
- * Stéphane Bresson, Centre des Technologies Nouvelles, Caen, tél : 02 31 46 25 26, bresson@ctn.asso.fr, <http://dnbn.ctn.asso.fr>

À lire également :

- * [1] PÉDAUQUE, R. T. Le document à la lumière du numérique, Caen, C&F éditions, 2006. 218 p.
- * [2] APROGED, La maîtrise du cycle de vie du document numérique, Paris, mai 2006. 32 p.

ENJEUX DE L'INNOVATION

Mais innover suppose aussi de prendre en considération toutes les conséquences possibles, qu'elles soient juridiques – la propriété intellectuelle devient l'élément essentiel de la capitalisation des entreprises –, culturelles avec une capacité de publication généralisée, organisationnelles, sociales et politiques.

Le document numérique est un enjeu majeur de la société de l'information. La Commission Européenne ne s'y est d'ailleurs pas trompée et parmi les sept défis du 7^e PCRDT, deux relèvent directement du document numérique : d'un côté le développement de l'intelligence dans les systèmes cognitifs et, de l'autre, le développement des bibliothèques numériques et des outils de création de contenus.