

Mémoire numérique : le droit à l'oubli

Le sujet du droit à l'oubli appliqué au web (puisqu'il s'agit principalement de cet espace numérique là depuis quelques années) a été traité sporadiquement, émergeant au gré d'incidents singuliers vécus par une personne ou un petit groupe (voir article « droit à l'oubli / googlerie » page 7 de ce document), de mouvements d'acteurs du numérique soulevant des interrogations quant à la portée de leur action sur la vie privée entre autres, ou encore lorsqu'un gouvernement fédéral exige des fournisseur d'accès Internet de lui livrer une année d'activité des internautes circulant via ses tuyaux pour l'aider à renforcer ses lois encadrant la pornographie (Etats-Unis).

Si la loi informatiques et liberté datant de 1978 a encadré en France l'application du droit à l'oubli dans une société glissant de l'informatisation à la virtualisation en passant par la numérisation, partout dans le monde des membres de réseaux sociaux, contributeurs de médias participatifs ou encore bloggeurs temporaires se débattent pour faire valoir ce droit devant la pertinence des machines à remonter le temps et les souvenir que sont les moteurs de recherche, le silence des commerçants de données personnelles (facebook par exemple) et enquêtes peu scrupuleuses d'un ministère du logement vérifiant sur des profils de réseaux sociaux la véracité des informations déclarés par les bénéficiaires d'allocations sociales (Canada). Exercer son droit à l'oubli à cet endroit, c'est pouvoir garder la main sur ses données personnelles, les éliminer ou faire disparaître au fur et à mesure ou exiger de ses interlocuteurs invisibles comme facebook d'effacer toute trace personnelle (ce qui peut demander du temps) ; en quelque sorte reprendre le contrôle sur la « machine », celle par laquelle, par ignorance ou faiblesse de jugement sur la portée de ses actions, nous nous déversons en grande majorité inlassablement, en toute insouciance.

Phénomène connexe, une forme de contradiction constituante se dresse aujourd'hui dans la montée en puissance du Web comme plateforme « première » des relations sociales, économiques ou politiques entre humains ; jamais l'intimité ne s'était dévoilée autant et avec autant de liberté ou d'impudeur (selon), suscitant notamment l'appétit (sur le web, subtilement, il est question de monétisation de l'audience) des adoreteurs d'une société régie par les lois de la mercatique (ce à quoi peu de gouvernements élus en occident voient quelque chose à redire). Par ailleurs, rarement la censure avait atteint de telles proportions, et immédiatement je relativise, dans sa dimension « visible, sans scrupule et décomplexée », moins dans une perspective historique. Aussi, pour prendre un seul exemple, exercer son droit à l'oubli lorsque vous êtes un blogueur chinois légèrement déviant (au regard du pouvoir), c'est contourner la grande muraille de feu en utilisant des systèmes comme Picidae, créé par deux artistes suisses et qui consiste en la transformation en images navigables non traçables de sites interdits. Un réseau dans le réseau, caché du réseau. Se faire oublier en quelque sorte, pour compenser l'absence de droit liée.

Par ailleurs, si la question du droit à l'oubli se pose de plus en plus sur le web, le point d'interrogation grandit aujourd'hui simultanément en dehors de l'écran (rétention de sureté, devoir de mémoire...)

Juste pour initier l'avant-débat.

Nous sommes partis dans l'idée, pour donner au débat (essayer du moins) une dimension participative de créer un décalage dans la manière d'aborder le sujet ; pas de théorisation ni de regard global (ce que je viens de faire), mais créer des questionnements-rebonds à partir d'une expérience personnelle, d'une interrogation liée à l'exercice d'une profession, d'un récit fictif...

Ci-dessous, quelques documents qui titilleront éventuellement votre réflexion. Rien d'absolument nécessaire, mais que vous pouvez compulser d'ici au 27 mars, si l'envie de picorer vous saisit :).

En attendant de vous voir, bonne journée et merci encore de votre participation.

Cordialement,

Laurent Vannini

1 - Dossier écrit dans le cadre d'un programme de veille sur les nouvelles technologies pour la R&D Orange. Sa thématique : la mémoire numérique.

Il ne traite pas spécifiquement du droit à l'oubli, même s'il l'évoque, mais permet d'aborder les dimensions multiples de la mémoire numérique (stockage, fonctionnalités, contenus...). Pour autant, bien des dimensions manquent encore (par exemple : qui accède à l'espace et outil de mémoire qu'est le web ? on peut parler de droit à la mémoire pour une grande partie de la population mondiale, 80% d'entre nous, qui n'a pas accès au web et ne peut s'y raconter...autre dimension zappée par manque de place dans ce document formaté, la dimension émotionnelle de la mémoire...la madeleine)

Je me permets de vous le glisser moins pour servir de base de réflexion au débat centré sur le droit à l'oubli, que pour éventuellement « augmenter » si besoin votre perception de ce que le champ de la mémoire numérique contient aujourd'hui (notamment en termes de services et technologies dédiés, qui ne sont pas forcément votre tasse de thé) et aider ensuite à circonscrire un périmètre pour votre intervention.

Homo numericus recherche souvenirs fragmentés

Septembre 2007, rédigé par Laurent Vannini

Sources :

Internet Actu - [Vers le web implicite](#) - 06/07/2007

Internet Actu - [Quelles recettes pour oublier dans un monde d'informatique omniprésente](#) - le 22/05/2007

ZDnet - [Watching cities in 4D](#) - 09/07/2007

Lorsqu'en 2004 le service de messagerie de Google voit le jour, les webmails gratuits du marché ont une capacité de stockage inférieure à 25 Mo, soit l'équivalent d'environ 7 chansons de 3 minutes compressées au format MP3. Aussi, le gigaoctet de données (soit 50 fois la capacité de stockage habituelle), évolutif de surcroît¹, proposé dès son lancement par « Gmail » paraissait invraisemblable ; plus la peine d'effacer un seul de nos messages, d'autant que ces derniers sont regroupés par conversation, facilitant ainsi classement et extraction. La « mémoire numérique » ne coûte plus rien...ou presque².

Bientôt, nous serons capables de « conserver » tout ce qui passe entre notre souris et nos synapses : ce que nous avons produit, produisons ou produirons; toutes ces données, visibles ou invisibles, privées ou publiques, implicites ou explicites qui ont un dénominateur commun, l'être humain, et constituent une sorte de mémoire numérique de l'activité humaine et « méta-humaine » sur l'internet.

Conséquence de l'installation progressive de cette « mémoire totale », l'humain déplace ses propres limites : désormais ce sont les fonctions mémorielles – traitement, extraction, pertinence, oubli – qu'il doit réinventer.

Les « mémoires » de l'être humain

La mémoire constitue avant tout la faculté de **stocker, conserver** et de **rappeler** des états de conscience passés, des expériences vécues ou des informations apprises. Elle structure notre vie et notre personnalité en stockant notre histoire, notre culture, nos actions. Les événements vécus s'y inscrivent avec ou sans effort, consciemment ou pas.

Dans un deuxième temps, la mémoire est apparentée au souvenir même, c'est-à-dire à ce qui remonte à la surface après remémoration individuelle ou collective. **La mémoire n'est jamais le fait, mais le souvenir du fait.** Il s'agit ici de ne pas confondre Mémoire et Histoire. Cette dernière propose, dans la distance avec l'évènement passé, une objectivation fondée non pas sur le souvenir d'une ou quelques personnes, mais sur la convergence d'un ensemble de « données » rationnelles.

Enfin dans un dernier grand ensemble, on associe la mémoire au support matériel de son « enregistrement » ou de sa conservation, c'est-à-dire aux outils qui au travers des siècles ont permis à

¹ Aujourd'hui, chaque utilisateur Gmail dispose d'environ 2.8 gigaoctets d'espace.

² Les utilisateurs Gmail peuvent désormais étendre la capacité de leur messagerie moyennant finance : 6 Go pour 20 \$ par an et jusqu'à 250 Go pour 500 \$ par an. Par ailleurs, Yahoo propose depuis mai 2007 une offre de stockage illimitée pour les utilisateurs de son service de messagerie.

l'homme, sciemment ou non, de fixer le récit de son histoire : du papier aux « molécules »³ en passant par l'analogique, les supports se miniaturisent et laissent à penser que chaque instant de la « réalité » du monde peut désormais être préservé et retrouvé.

La baisse continue du prix du Giga-octet, le développement de nouvelles technologies proposant de meilleures densités de stockage ou encore la simplification des interfaces logicielles et procédures de stockage (généralisation de la pratique du « glisser-déposer ») permettent d'envisager à court terme la possibilité de « tout » numériser.

Les supports et outils numériques « ex-cortex » de la mémoire

Créer une Intelligence Artificielle – une machine douée de capacités logiques, réflexives, d'analyse identiques à celles de l'homme et dotée de surcroît d'une mémoire infailible et totale – est depuis longtemps un « rêve/cauchemar » humain, porté entre autres par les écrivains de Science Fiction.

Côté « machine », des chercheurs d'IBM ont publié en mai 2007 deux rapports portant sur des simulations à grande échelle de réseaux de neurones. L'objectif était de reproduire un système proche du cerveau d'une souris, tant par la taille que par le nombre de connexions et leur activité. Soit le fonctionnement de 16 millions de neurones avec 8 000 synapses par neurones. Ces travaux ne reproduisent pas le fonctionnement d'un cerveau, mais mettent en évidence la faisabilité de simuler un nombre important de connexions et d'étudier l'échange d'informations à cette échelle. Côté « humain » cette fois, la mémoire fait l'objet d'une recherche permanente afin que les supports neuronaux conservent leur « plasticité » originelle : une équipe de l'Université McGill (Montréal) a découvert un gène mutant améliorant la mémoire à long terme chez la souris de laboratoire. Prochaine étape : un médicament destiné à améliorer la mémoire chez l'humain, pour redonner une qualité de vie aux patients atteints de la maladie d'Alzheimer par exemple.

Plus simplement, d'autres découvertes récentes ouvrent les portes vers une mémoire numérique de plus en plus performante : on peut citer l'utilisation des matériaux à changement de phase, déjà utilisés pour le stockage de données optiques (DVD réinscriptibles) ou encore des Mems (système mécanique sans transistor dans lequel les bits d'information sont portés par des interrupteurs mécaniques et microscopiques). Et pour l'extraction des données, le stockage en mémoire holographique (mémoire dense par stockage de données dans des cristaux ou des polymères photosensibles) offre des résultats étonnants : une équipe de chercheurs au Japon a développé un système de reconnaissance faciale très rapide à l'aide d'un corrélateur optique, qui associée à une base de données d'images stockées sur mémoire holographique, permet d'identifier 1 million de personnes par seconde. L'équipe assure avoisiner 99% de taux de réussite à l'extraction d'un visage d'une foule sur photographie. Tendre vers une mémoire numérique « infailible » peut sembler une réponse appropriée à une angoisse grandissante au sein de nos sociétés de l'information, l'angoisse de perdre la mémoire.

Travailler sa mémoire : le marché des seniors et des « apprenants »

Exercices visuels, calcul mental, méthodes d'ancrage des souvenirs dans le corps, pensée positive, méthodes respiratoires ; des centaines de techniques sont proposées en ligne au commun des mortels pour « guérir de ses amnésies » et mieux (ou plus) se souvenir. [Happy Neuron](#) propose une série de quarante exercices avec coach pour améliorer l'ensemble de ses capacités cognitives (langage, mémoire, logique...). Sorti au mois de juin 2006, *Programme d'Entraînement Cérébral du Dr Kawashima : quel âge a votre cerveau* disponible sur Nintendo DS a été acheté par dix millions de personnes à travers le monde.

Si les personnes âgées sont les premières touchées par des déficiences mémorielles, tous se posent au moins à un moment de leur vie la question de leur « mémoire ». Combien de personnes rencontrons nous chaque jour, se plaignant de ne pas avoir de « mémoire » ? Une étude menée au Trinity College de Dublin au printemps 2007 montre que les moins de trente ans, ceux-là même qui utilisent quotidiennement depuis leur enfance ordinateurs, téléphones, Palms et autres « béquilles » technologiques, externalisant ainsi leur mémoire, sont ceux qui la « perdent » le plus...

³ Des chercheurs de deux instituts californiens ont présenté dans la revue *Nature* du 25 janvier 2007 un prototype de mémoire d'une capacité de 160 Kb et de la taille d'un globule blanc, ce qui correspond à la première mémoire d'une telle densité encore jamais fabriquée.

Soutenir et assister sa mémoire

Outils d'archivage, de classement ou de recherche, de gestion des flux d'information ou encore aide-mémoire forment « l'arsenal » de la mémorisation numérique individuelle. De l'organisation des favoris dans notre navigateur à la réception des informations thématiques issues de centaines de sites présélectionnés, notre mémoire d'utilisateur du web est épaulée en permanence. L'une de ces béquilles, [Hanzoweb](#) est un outil social d'archivage du web comparable à Del.icio.us dans le sens où il permet d'archiver et de tagger ses sites favoris. Il ajoute une composante de mémorisation « physique » des pages web sélectionnées qui peuvent alors à tout moment être visionnées sur un compte personnel disposant de 100 Mégabyte par mois. La *folksonomy*, les outils de *tagging* ou encore les outils de recherche sémantique⁴ ajoutent une fonctionnalité d'archivage et de recherche proche de la manière dont notre cerveau fonctionne pour garder en mémoire des faits ou informations.

[MyLifeBits](#)

Avec 2500 Go de données par année, les équipes de Gordon Bell ont été confrontées au problème de surcharge informationnelle. Pour fournir une interface intelligente capable de stocker toutes nos données personnelles, mais surtout de les retrouver avec rapidité et flexibilité, un des choix technologiques effectués par les équipes de MyLifeBits a été de ne pas créer d'arborescence d'archivage de données mais d'offrir plusieurs modes de visualisation des mêmes données sur un axe chronologique représentant le fil de la vie, par type de média, ou par collections thématiques.

Laisser ses traces ou retrouver celles de ses proches

Si les blogs, vlogs ou autres mlogs ont marqué le point de départ d'une tendance de plus en plus profonde à l'expression subjective d'une portion de sa vie sur le web ou sur mobile, qu'elle soit de l'ordre de l'intime ou du professionnel, ils ont été rapidement suivis par le phénomène du LifeStream, consistant à agréger chronologiquement en un seul endroit ses différentes activités ou productions en ligne ou hors ligne ; Del.icio.us, Flickr, Twitter, Plaxo Pulse sont autant d'outils permettant à tout moment à l'utilisateur de donner à voir où il en est de son cheminement personnel, mais également de partager avec d'autres sa mémoire ; l'addition de ces mémoires partagées et leurs interconnexions constituent des réseaux assimilables à une mémoire sociale du web dont les passerelles et portes d'entrées ne sont autre chose que les tags, les métadonnées...

Les LifeLogs⁵ symbolisent l'étape suivante consistant à systématiser l'enregistrement et l'archivage de l'ensemble de ces données de vie. A l'instar de Hewlett-Packard ou du MIT qui ont initié des travaux relatifs aux LifeLogs, une équipe de recherche s'est constituée chez Microsoft autour du projet [MyLifeBits](#), véritable base de données historique et personnelle. MyLifeBits requiert l'utilisation d'un système portable de capture, la SenseCam, sorte de boîte noire personnelle programmée pour effectuer jusqu'à 2000 photos par jour. Depuis 2001, Gordon Bell, un des membres de ce groupe projet, s'est constitué son lifelog. Il a scanné, capté, numérisé et stocké l'ensemble des informations – photos, livres, pages web, films, sons – auxquels il a eu accès, ainsi que l'ensemble des traces de ses déplacements, de ses échanges et communications. Au-delà de ces contenus, un capteur personnel permet également de mémoriser des choses dont nous ne sommes même pas conscients, comme les taux d'oxygène dans le sang, les taux de CO2 dans l'air, le rythme des battements de cœur et autres constantes physiologiques.

Autre dimension de l'individu à être passée au crible, la famille. Le début du XXI^e siècle a vu un grand nombre de sites et logiciels de généalogie se développer. D'après une étude Ipsos pour Notrefamille.com, une famille sur deux possède déjà un arbre généalogique et Internet est devenu la principale source d'informations dans ce domaine. L'un des plus utilisés en France, Genealogie.com a récemment vécu une évolution vers le web 2.0, offrant des services communautaires, la possibilité de créer son arbre généalogique en ligne ainsi qu'un outil de géolocalisation.

Tout ou presque de l'individu, de son identité, de sa mémoire se trouve entreposé publiquement ; même parfois ce qu'il n'avait pas l'intention de déposer, mais qui possède une grande valeur aux yeux des acteurs marchands de l'économie numérique.

⁴Lire l'article SD « [Freebase : Web sémantique et Web2.0 s'unissent pour déchiffrer le monde](#) »

⁵ Système de stockage qui enregistre automatiquement et systématiquement, puis archive tout ou partie des dimensions informationnelles de vie d'un objet ou d'un individu.

La mémoire implicite de l'homme numérique

Une économie de la « mémoire » implicite du web se développe trouvant ses origines dans le système de recommandation d'Amazon⁶. Lorsque nous aimons un article, nous passons du temps à le lire, quand un film nous plaît ou un morceau de musique, nous le recommandons à nos amis. Nous agissons ainsi implicitement. Ce « web implicite » met en évidence le cheminement inconscient mais structurant de notre parcours « internautes », constituant une forme de mémoire passée et projetée. Un exemple avec [Last.fm](#), moteur de recommandation de musique qui se base sur votre bibliothèque d'artistes préférés en ligne pour vous recommander des chansons que vous ne connaissez pas. « Nous sommes donc passés d'une toute puissance du lien hypertexte, point nécessairement nodal de développement du réseau et des services et outils associés, à une toute puissance du "parcours", de la navigation "qui fait sens", de la navigation "orientée" au double sens du terme », explique sur son blog [Olivier Ertzscheid](#), maître de conférences en sciences de l'information. Il parle de troisième révolution du web, lorsque les moteurs, sans requête de notre part, sauront selon l'heure de la journée, le lieu de connexion, l'environnement, ce que nous sommes susceptibles de chercher. Il met par ailleurs en garde contre la monétisation systématique de ces parcours.

La mémoire comme outil marketing du publiciste numérique

Ce que nous mémorisons, tant sur le web que sur d'autres écrans d'affichage publicitaire, intrigue les marketeurs. [Imagiin](#) est une agence de publicité proposant des services de mesure de l'efficacité des publicités à base de taux de mémorisation des annonces. Des volontaires visionnent un certain nombre d'écrans publicitaires en ligne et sont rémunérés lorsqu'ils répondent à trois questions portant sur les annonces vues. Dans leur communiqué de presse, Imagiin souligne sa « promesse de transformer l'opportunité d'être vu (ODV) en certitude d'être vu, compris et mémorisé (VCM) ».

TF1 de son côté travaille avec [Impact Mémoire](#), « cabinet conseil en efficacité mémorielle », dont le rôle est d'explorer grâce à l'Imagerie à Résonance Magnétique (IRM), technique jusqu'alors réservée au domaine hospitalier, ce que le cerveau retient dans la communication publicitaire. Les « neuromarketers » ont constaté que la zone du cerveau réactive aux images publicitaires est associée à l'image de soi. Il convient donc d'activer au maximum ce cortex pour transformer tout amour de soi en tant que soi – le narcissisme – en amour de soi en tant qu'autre – une cible publicitaire. De la mesure de la mémoire à l'influence. Ce qui n'est pas sans rappeler les images subliminales injectées dans des films de campagnes électorales pour inciter à un certain vote.

Mémoires collectives numériques, entre reconstruction et reconstitution de l'histoire

La mémoire collective est, selon l'historien Pierre Nora « le souvenir ou l'ensemble de souvenirs, conscients ou non, d'une expérience vécue et/ou mythifiée par une collectivité vivante de l'identité de laquelle le sentiment du passé fait partie intégrante ». Sur le Web, cette rencontre entre psyché et social se fait de plus en plus souvent sur les espaces collaboratifs de type wiki, forums ou blogs ou chaque personne est amenée à apporter sa pierre à un édifice conversationnel, thématique, artistique ou autre. Le site [memoryarchive.org](#), **agrégation** sur un wiki de **souvenirs d'internautes concernant des événements historiques**, représente cette première forme numérique de mémoire collective. Dans le même ordre d'idées, juste après la fusillade sur le campus de Virginia Tech, le 16 avril 2007, une page [Wikipédia](#) était créée et devenait la source d'information de milliers d'internautes et des plus grands médias ; en 24 heures à peine, 2000 contributeurs avaient participé à la rédaction de cette page, réalisant environ 9 000 modifications. Cette forme de reconstruction d'un événement par **agrégation de mémoires individuelles** tend à fondre en une seule entité, « l'information » et l'événement en question, tout en omettant une dimension historique, c'est-à-dire le temps de travail de l'historien sur un événement dans un contexte spatio-temporel donné (prise en compte de variables sociales, sociétales, économiques...). La masse d'informations qui nous est livrée chaque seconde rend le présent « permanent et immanent » et ne nous permet pas toujours de prendre le recul suffisant pour analyser une information issue d'un fait et dissocier cette information du fait réel. Pour autant, l'agrégation de mémoires individuelles peut constituer un vivier de documents pour la reconstitution cette fois, et non la reconstruction, d'une histoire collective, à commencer par les éléments au plus proche de chacun d'entre nous, ceux de notre territoire de vie.

⁶Amazon utilise un système de recommandations de produits à ses clients sur la base de 4 éléments : les produits déjà achetés par le client, les produits que le client dit déjà avoir, les produits consultés par le client, les liens entre les produits, extrapolés à partir des listes d'achats, créées par l'ensemble des clients.

La mémoire sociale prend racine dans un territoire

Caen, Marseille, le Finistère ou encore des villages comme Le Faou et ses 1500 âmes ont en commun de recueillir la parole (orale, écrite ou imagée) d'habitants pour écrire la mémoire locale de leur territoire. L'association arkae par exemple œuvre pour la conservation de la mémoire des habitants de la commune d'Ergué-Gabéric, comptant 7 000 habitants. Depuis trente ans, elle collecte les témoignages et a créé un [wiki](#) pour archiver et rendre public son travail ; On y trouve scans de documents datant du 15^{ème} siècle, historique de l'orgue trônant dans l'église du bourg ou encore histoires de sorcellerie racontées par une villageoise. Rues et lieux-dits sont scrupuleusement détaillés donnant à voir une évolution historique du lieu, ancrée dans une réalité sociale mouvante.

Cette dimension évolutive d'un ville au fil du temps a plus récemment fait l'objet d'une représentation en « 4 dimensions », faisant apparaître l'historique des modifications urbaines. Les chercheurs de l'Institut de technologie de Géorgie et de Microsoft Research ont scanné de nombreuses photos de la ville d'Atlanta, prises approximativement du même endroit. Le logiciel développé, nommé [4D cities](#), identifie ensuite les structures en 3D dans l'image et les transforme en séries de points, réorganise chronologiquement les photos, autorisant les chercheurs à construire et animer une vue 3D de la ville dans laquelle il est possible de voyager « temporellement ».

Reconstituer, préserver et transmettre Histoire et Patrimoine

Il existe par ailleurs une multitude de plateformes de recueil et conservation de témoignages, d'archives nationales, d'articles concernant une période donnée de l'histoire, une communauté ou une thématique. Cela peut être le fruit d'un travail individuel dans [parolesdesclavage.com](#), blog consacré au recueil de témoignages des derniers esclaves de Martinique qui donnent à entendre au travers de leur vécu une histoire ancrée dans notre inconscient collectif ; ou la commercialisation d'une base de données de dimension internationale comme [ancestry.com](#), site de généalogie états-unien, qui a récemment ajouté à sa base de données 7,5 millions de noms et données identitaires issus des recensements d'indiens d'Amérique du nord, soit 150 années d'histoire de populations natives issues de 250 tribus vivant dans 275 réserves.

Au sein de l'entreprise, concept et outils de Knowledge management apparus dans les années 80 pour répondre à un besoin de pérennisation et de tri de l'information ont vu de nouvelles formes de « captation de la mémoire des savoirs et savoir faire » naître de l'utilisation des outils du web 2.0, et plus spécifiquement des outils de Mind Mapping⁷. Ainsi [The Brain](#) est une plateforme logicielle qui permet aux entreprises d'organiser et gérer leurs informations et données de manière visuelle, et non hiérarchique.

Dans l'espace public cette fois, à un niveau européen, le Centre Virtuel de la Connaissance sur l'Europe (CVCE) a développé en 2004 l'[European Navigator](#), plateforme éducative accessible en 5 langues par une interface en format flash, qui propose des documents vidéo, audio, textuels ou sous la forme de jeux sur l'histoire de l'Europe et de ses institutions depuis 1945, permettant à l'utilisateur de se construire son opinion.

D'autres sites de « mémoire partagée », à caractère professionnel comme le projet [archives ouvertes](#) en France (mise à disposition de données scientifiques entre pays pour permettre un travail pluridisciplinaire et transfrontières) ou grand public, comme [Open Library](#) (voir encadré) participent de la construction d'une mémoire collective « globale » et rejoignent une volonté transnationale de sauvegarder des archives aujourd'hui disponibles sur des supports périssables.

Bibliothèques numériques

Depuis le lancement de Google Books (ex Google Print) en 2004, la numérisation et mise à disposition d'écrits (du domaine public particulièrement) ont fait l'objet de nombreux projets publics ou privés. Il faut distinguer les bibliothèques qui mettent en ligne des images numérisées des pages des livres originaux ([Gallica](#), bibliothèque numérique de la BNF qui présente 90 000 livres) et celles qui mettent en ligne le texte lui-même (comme [Gutenberg](#), environ 17000 livres disponibles).

Dernier projet en date, [Open Library](#), propose de rassembler toute l'écologie informationnelle de l'univers des livres sur une plateforme wiki. Pour les livres non diffusés sous licences libres, ou pas encore dans le domaine public, Open Library propose d'accéder à des ressources documentaires permettant de s'informer sur ces ouvrages et de les annoter.

⁷ Une carte heuristique (mind map en anglais), aussi appelée carte des idées ou carte conceptuelle est un diagramme qui représente les connexions sémantiques entre différentes idées, les liens hiérarchiques entre différents concepts intellectuels. À la base, il s'agit d'une représentation principalement arborescente des données.

En France, l'INA doit devancer la dégradation des supports de type analogiques. L'institut a atteint 50% de son objectif de sauvegarde d'ici à 2015 des 835 000 heures d'archives TV et radio menacées et propose déjà 250 000 heures de programmes aux professionnels et grand public sur le web.

Le web, mémoire totale et oubli permanent

La mémoire volontaire ou explicite du web

Avec plus de 110 millions de sites web répertoriés, le World Wide Web peut être comparé à un gigantesque « inconscient collectif » numérique, dont il est parfois difficile d'extirper ce que nous cherchons et dont la « pertinence » pour tous n'est pas acquise. Est-ce que le Web doit alors faire l'objet d'un archivage systématique global ? Si une sélection doit s'opérer, sur quels critères ? Et par qui ? L'équipe d'Archive.org, site états-unien non marchand, répond partiellement à cette question depuis 1996. Elle répertorie pages web, mais aussi textes, films, enregistrements de concert, et plus généralement tout type de documents, pourvu qu'ils soient numérisés. Les archives ainsi constituées se comptent en milliards. Au-delà de sa fonction de bibliothèque des contenus numériques, Archive.org permet d'appréhender l'évolution du web en termes de contenus, d'ergonomie, de design.

En France, l'INA et la BNF ont, depuis juin 2006 et le vote de la loi DADVSI⁸, la charge du dépôt légal⁹ du web. Concrètement, l'INA collecte les sites relevant de la communication audiovisuelle, et la BNF tous les autres. La seule photo instantanée du domaine .fr réalisée en début d'année 2005, représentait un volume de 27 téraoctets alors qu'à cette date, les plateformes d'UGC (comme daily motion ou flickr) n'existaient pas encore.

Pas étonnant alors que pour trouver son chemin vers le contenu souhaité, le « souvenir numérique », l'être humain ait besoin de « béquilles ». Que devient alors la fonction de mémoire humaine si elle est remplacée progressivement par une fonction de recherche de l'information dans une mémoire collective numérique accessible à tout moment ? Est-ce qu'en externalisant ainsi la fonction de mémorisation, on ne perd pas progressivement la fonction humaine imparfaite de « mémoire », et par la même occasion une partie de son identité ; en nous référant aux mêmes sources, émotions et ?? les personnalités sont gommées, tout devient lisse, aseptisé.

Si nous développons des facultés nouvelles (navigation et recherche sur internet, traitement multitâches...), nous devons être connectés pour garder la trace d'informations importantes : que se passe-t-il lorsque nos outils tombent en panne ? Ou plus important, lorsque chaque accès à nos espaces de données de vie devient payant ?

L'oubli comme moteur de la mémoire

L'être humain se trouve confronté à un monde numérique où tout est mémorisé, rien ne peut s'oublier. [Christophe Deschamps](#), du blog « outils froids » s'interroge sur les changements sociaux générés par une situation inédite où l'oubli devient impossible. « *Tous les travaux récents portant sur le cerveau montrent que pour vivre l'homme a autant besoin de se souvenir que d'oublier et il en va de même pour la société.* » Et même s'il est fondamental de se souvenir en premier lieu pour pouvoir ensuite oublier, la résurgence permanente d'éléments liés à une situation qu'une personne, ou société, est parvenue à « oublier », à « transformer », fonctionne comme un boulet attaché à sa cheville. Ce que nous produisons sur le web à une époque est facilement extractible dans le présent ou le futur, ne permettant pas l'oubli de ce qui a été dit, fait, pensé mais venant buter sur le présent comme pour rappeler les fautes, erreurs ou différences et ne pas permettre le « changement ». Or, notre identité, qui se construit tant sur notre mémoire que sur nos oublis, est en mouvement, autant que peuvent l'être nos pensées et nos souvenirs. Le « deuil » même devient difficile avec des sites comme MyDeathSpace.com, où sont répertoriés noms et causes de la mort des membres décédés de MySpace.

Un exemple à méditer avant de continuer à « tagger », « bookmarker », enregistrer, sauvegarder et nous plaindre de notre incapacité à mémoriser. Le blog de Lewis Trondheim, [les petits riens](#), est un carnet de bord dessiné proposant régulièrement une anecdote de la vie de l'artiste sous forme de planche. Les planches les plus anciennes s'effacent progressivement pour disparaître complètement après que cinq

⁸ « Droit d'Auteur et Droits Voisins dans la Société de l'Information » Loi française issue de la transposition en droit français de la directive européenne 2001/29/CE sur l'harmonisation de certains aspects du droit d'auteur et des droits voisins dans la société de l'information

⁹ Le dépôt légal, mis en œuvre sous François I^{er}, est l'obligation faite à tout éditeur, imprimeur, producteur, distributeur, importateur de documents d'en effectuer un dépôt auprès des organismes désignés par la loi.

nouveaux épisode ont été publiés. Au-delà de l'humour et de l'autodérision émanant de l'œuvre, des planches « évanescentes » qui donnent à l'œuvre un caractère éphémère et nous poussent à célébrer la valeur du temps présent consacré à son observation, sa lecture, ses émotions puisque « plus jamais », en l'état, la même expérience ne pourra se reproduire.

.....

2 - Quelques articles ou textes, empruntés au Web et leurs auteur-es bien réels, traitant de la question du droit à l'oubli dans un contexte de numérisation totale.

Quelles recettes pour oublier dans un monde d'informatique omniprésente ? www.internetactu.net

Le 22/05/2007, par [Hubert Guillaud](#)

[Viktor Mayer-Schönberger](#), professeur de droit public à l'université d'Harvard, s'inquiète de la capacité de mémoire de notre environnement. "Comme l'Etat soviétique, Google n'oublie rien", explique avec provocation le professeur. Dans un [récent article](#) il note que l'expansion de la capacité de stockage de l'informatique induit que de plus en plus d'éléments de nos vies vont être enregistrés et sauvegardés. Cela va profondément influencer notre manière de voir le monde et notre comportement. Que se passe-t-il si ce que nous faisons peut-être utilisé contre nous des années plus tard, si tous nos commentaires impulsifs sont préservés ? L'absence d'oubli peut nous conduire à une société où nous parlerons moins librement et moins ouvertement...

Mayer-Schönberger propose une solution : que la mémoire de nos machines et logiciels se contraigne à fonctionner comme elle l'a fait depuis des millénaires, et que l'on repasse du souvenir perpétuel à l'oubli avec le temps. La loi devrait prévoir, pour différents types d'outils ou de logiciels, des durées après lesquelles l'information doit disparaître : quelques jours ou quelques semaines pour les caméras de surveillance, peut-être quelques années pour l'enregistrement de nos achats chez Amazon (après tout, nos préférences littéraires doivent-elles être constantes pour l'éternité ? Comment continuer à les faire évoluer d'ailleurs, si les systèmes continuent de prendre en compte le poids des données passées ?). Les utilisateurs devraient même pouvoir changer eux-même la date d'expiration des données qu'ils conservent ou laissent conserver par d'autres, même si par défaut, l'oubli devrait être la norme.

Le but de Mayer-Schönberger : réduire la masse de données que gouvernements et entreprises savent collecter sur les utilisateurs et citoyens, simplement parce qu'ils le peuvent. Mettre une date d'expiration par défaut forcera ces organisations à s'interroger pour savoir si elles ont vraiment besoin d'enregistrer ces données pour toujours. Nous aussi.

.....

Le deuxième article écrit en 2003 est intéressant à la fois pour son contenu, mais également pour les nombreux commentaires l'accompagnant et qui donnent à entendre des voix différentes sur le droit à l'oubli dans un monde « google-isé ».

Le droit à l'oubli / googleries <http://www.davduf.net>

Publié: samedi 19 avril 2003 - Auteur(s) : davduf

le net est une belle machine diabolique. Ici règne la mémoire, ni tout à fait morte, ni jamais vraiment vive. Le net se souvient de tout. Du meilleur comme des erreurs. Des milliards de pages, autant de traces indélébiles. Cas d'école : les décisions de justice rendues publiques. A jamais, pour toujours. Comme un flicage anodin.

C'était il y a quelques mois. Un procès parmi d'autres. Dans le box, Yannick [\[1\]](#), jeune homme accusé d'avoir blessé sa petite amie à coups de couteau. Sur les bancs de la presse, *Libération* et d'autres. Les témoins défilent, l'audience est courte, le jugement tombe. Tentative d'homicide volontaire, six ans de

prison pour Yannick. Et l'affaire s'arrête là. Les parties civiles sont indemnisées, le condamné purge sa peine, et tous sombrent dans l'oubli. Quand, la semaine dernière, un email parvient à *Libération*. C'est Yannick. Après deux ans et demi de détention préventive, et six mois de prison, il est libre. Sous liberté conditionnelle jusqu'en 2002.

Mais voilà. Yannick est confronté à un « *gros problème* ». Le compte-rendu du procès, publié par *Libé*, figure toujours, plusieurs mois plus tard, sur le Web du journal, en accès libre. D'où son email, pour dire : « *si quelqu'un tape mon nom et mon prénom [dans un moteur de recherches de type voila.fr, nomade.fr ou yahoo], il va tomber sur cette triste affaire. De là je risque de perdre mon emploi, je crois qu'il est facile de comprendre pourquoi, et si je perds mon travail, la chance de réinsertion que l'on m'aura donnée n'aura servi à rien puisque je devrai retourner terminer ma peine en détention...* »

« On a déjà un casier judiciaire, on a pas forcément envie d'un avoir un autre »

Joint par téléphone, Yannick en dit plus. Devenu formateur à l'Internet, il parle de son « *appréhension* » devant ses stagiaires, quand il leur explique le fonctionnement des moteurs de recherches. De la « *peur* » que ses élèves fassent comme tout néophyte : taper leur nom, puis ceux de leurs proches, puis, tiens, pourquoi pas, celui de leur formateur... Dans la voix de Yannick, comme dans son courrier, ni colère ni rancœur mais beaucoup de remords pour « *ce qui s'est passé* » et un profond respect pour « *le droit à l'information* ». Seulement, lui qui « *aspire à une vie nouvelle* », qui se « *réadapte petit à petit à la vie en société* », qui « *a payé sa dette* », demande que l'article en ligne soit expurgé de son nom de famille. Une requête au nom du droit à « *certaine tranquillité* ». *Libé* s'exécute. Le problème, plus général, reste, lui, entier. « *On a déjà un casier judiciaire, on a pas forcément envie d'un avoir un autre* » lâche Yannick.

C'est le revers de la puissance de l'Internet. Les machines n'oublient rien. Avec le Net, la mémoire se fait mondiale, intacte, et directement accessible de partout, et par n'importe qui. Un événement qui fait l'objet d'une publication, un jour, dans un quotidien national, ou ailleurs, devient immédiatement sur le réseau une archive numérique, propice à tous les recoupements, à tous les traitements automatisés, et, donc, à tous les détournements d'usage. L'affaire Yannick est un cas d'école, parmi d'autres. Elle est reproductible à l'infini. Avec elle, c'est la question du droit à l'oubli qui surgit, de plein fouet. Une notion qui existe dans la philosophie générale de la loi française, avec son lot d'amnesties, de réhabilitations et de peines suritaires effacées des casiers judiciaires après cinq années. La jurisprudence est même très nette sur ce point : « *l'actualité crée une exception vis à vis de l'atteinte à la vie privée, explique Thierry Lévy, avocat au barreau de Paris. Le droit à l'information joue à plein [2]. Mais une fois que le fait divers est retombé, quiconque peut faire valoir le droit à l'oubli. Il revient alors aux tribunaux d'en apprécier le fondement* ».

Jusqu'ici, il était en fait souvent inutile d'en passer par là. Le droit à l'oubli s'appliquait de facto, une information quotidienne - télé, radio ou presse écrite - chassant l'autre. Mais l'Internet, lui, change la donne. Sauf intervention express, il rend tout imprescriptible, ou presque. Un changement d'échelle vertigineux, qui prend tout le monde de court, des éditeurs au moindre webmestre, tenté par la publication de n'importe quel type d'information.

« L'anonymisation » totale des décisions de justice

Conscientes du danger, plusieurs instances se sont pourtant déjà penchées sur la question. Ainsi, le Ministère de la Justice, qui a imposé, dans le cahier des charges de la diffusion sur Internet des données jurisprudentielles, « *l'anonymisation* » totale des décisions de justice. Ainsi du gouvernement qui, avant de basculer le Journal officiel sur l'Internet, avait saisi il y a peu la Commission nationale de l'informatique et des libertés (CNIL). Pour voir si la publication, à l'échelle mondiale, de certaines données n'allait pas entraîner quelques atteintes graves à la vie privée. La CNIL en trouva une, apparemment anodine : la liste des naturalisations. Ce qui semblait légitime en France pouvait, en effet, faire courir le risque à certains ressortissants ayant abandonné leur nationalité d'éventuelles représailles sur eux, ou leurs proches, dans leur pays d'origine. Etc.

N'empêche. Pour l'heure, l'expectative prévaut. Aux Etats-Unis, bon nombre d'employeurs exploitent, déjà, le filon Internet comme source d'informations personnelles, et pistent toutes les traces laissées sur le réseau avant de faire leur choix. En Europe, ici ou là, les avis se multiplient sans grande harmonie, et

parfois, les recommandations se superposent. Signe du flou qui règne : le rapport sur « *L'information émanant du secteur public* » du groupe des autorités en charge de la protection des données dans les Etats membres de l'Union européenne [3]. Le groupe ne peut que se borner à recommander aux pays de la Commission européenne de s'en remettre à « *l'appréciation au cas par cas [pour savoir] si une donnée à caractère personnel peut être publiée/accessible ou non, et si oui, dans quelles conditions et sur quel support* ».

dans les semaines qui ont suivi la publication de cet article (paru à l'origine dans *Libération*, 21 avril 2000), Emmanuel Lesueur de Givry, Conseiller à la Cour de cassation et Directeur du Service de documentation et d'études d'icelle, notait dans son rapport sur [La question de l'anonymisation des décisions de justice](#) :

*« (...) la possibilité offerte par le réseau internet et les performances des moteurs de recherche d'accéder en permanence et pour une durée illimitée aux noms des parties est sans commune mesure avec la publicité de la décision de justice ; alors que le casier judiciaire national fait l'objet de dispositions législatives protectrices, le risque de constitution de fichiers civils de locataires ou d'emprunteurs défaillants, de salariés ayant eu recours à la justice prud'homale, de mauvais conducteurs etc., n'est pas hypothèse d'école. Sommes-nous prêts, comme mentionné dans le »rapport Mandelkern« , à accepter la création de sites qui, dans certains États américains, permettent de »savoir à tout instant qui est en prison et pour quelle durée« , la localisation des »personnes condamnées pour délit sexuel, un plan des villes permettant de déterminer avec plus de précision leur domicile ?« Est-il acceptable que des élèves qui tapent les noms et prénoms de leur professeur dans un moteur de recherche découvrent que celui-ci a été condamné pour avoir blessé sa petite amie alors qu'il »a payé sa dette ? « (voir à ce sujet l'article paru dans le journal *Libération* du 11 avril 2000). Le détournement de finalité des fichiers de jurisprudence n'ouvre t-il pas la voie à une sorte d'affichage permanent, négation du droit à l'oubli, sorte de peine complémentaire perpétuelle ! ».*

[1] Pour des raisons d'anonymat, le prénom a été changé

[2] En France, la loi autorise la mention nominative de toutes personnes majeures impliquées dans une affaire judiciaire.

[3] 3 mai 1999

A propos de cet article

Forum

[Répondre à cet article](#)

1. 1 Ph R 19 avril 2003 @ 22:09

> le droit à l'oubli | googleries

Intéressant débat. Nous avons eu le cas sur www.monde-diplomatique.fr, d'une personne ayant signé une pétition en 1978 - pétition republiée 20 ans après sur notre site. Une recherche Google sur son nom faisait ressortir « en premier » cette pétition, et non ses travaux universitaires... fallait-il supprimer son nom de la pétition, comme elle nous le demandait ? La conséquence aurait été de « falsifier » quelque peu la pétition en question... pris entre ces deux exigences, celle de la vérité historique et celle du « droit à disparaître de Google », nous avons coupé la poire en deux (avec son accord), en ne maintenant ses initiales.

[Répondre à ce message](#)

2. 2 Meryem Marzouki 20 avril 2003 @ 12:33

> le droit à l'oubli | googleries

Oui, intéressant débat, comme tous ceux qui portent sur des conflits de droits aussi légitimes les uns que les autres. Ce n'est d'ailleurs pas spécifique à Internet, cf. le débat houleux sur le droit à l'information vs. le droit à la protection de l'image et de la vie privée, notamment lors de l'adoption de la loi Guigou sur la présomption d'innocence et le droit des victimes. La presse écrite était alors montée au créneau.

Pour revenir à nos pétitions, nous avons eu les mêmes demandes de retrait de signatures individuelles avec les pétitions lancées par [IRIS](#), et spécialement avec une [pétition contre MicroSoft](#), une action menée en 1998. Au début, nous laissions la liste de toutes les signatures individuelles en lecture publique, comme n'importe quelle page du site. Mais certains signataires, jeunes fous à l'époque, se sont retrouvés sur le marché du travail informatique quelques années plus tard, et ont compris leur douleur lors des entretiens d'embauche. Ils nous ont donc demandé la suppression de leurs signatures.

Après des débats internes assez chauds (« le droit à l'information », voire « la vérité historique », mais aussi « le droit à l'oubli » et « on ne nous demande quand même pas d'effacer les crimes de Staline », je vous passe les détails), nous avons convenu, pour toutes nos pétitions, d'afficher sur le site la liste des signatures individuelles pendant toute la durée d'une campagne de pétition puis (car il faut aussi savoir clôturer une pétition !) de les archiver en ne laissant à disposition publique que le nombre de signatures. Les signataires sont [prévenus](#) au moment de signer.

Ainsi, nous avons le sentiment de respecter à la fois le droit à l'oubli des personnes et le droit à l'information de tous. Cela nous semble une bonne solution, au-delà de nos seules pétitions : en effet, si une recherche scientifique doit être menée, par exemple, l'information reste disponible, bien que non publique. On peut aussi imaginer la situation où une personne se prétendrait diffamée si on disait d'elle qu'elle avait signé un texte ; en cas de procès, les archives pourraient prouver que ce n'est pas le cas.

Cette solution a d'ailleurs été adoptée pour les pétitions lancées par d'autres associations et syndicats du [R@S](#), qui utilisent le même système de collecte et gestion de signatures. Voir par exemple la dernière campagne en date, portant sur la [sauvegarde du droit d'auteur dans la création salariée](#).

Bien entendu, cela ne s'applique qu'aux signatures individuelles. Les signatures de collectifs restent, elles, toujours à disposition publique.

[Répondre à ce message](#)

3. *3 21 avril 2003 @ 12:25*

> le droit à l'oubli | googleries

Bonjour,

N'est-ce pas ce même mensuel qui, il ya de nombreuses années, évoquait que, aux USA, les informations personnelles des registres d'état civil, fiscaux, judiciaires etc. étaient publiques (seul leur éparpillement limitant leur prospection), et disponible depuis tout le territoire depuis les numérisations puis raccordements aux réseaux informatiques de ces données ?

[Internet, ou la fin de la vie privée](#)

[Répondre à ce message](#)

4. *4 Thomas 30 avril 2003 @ 14:52*

> le droit à l'oubli | googleries

L'article est intéressant et émouvant. Ironie du sort dans le fait que ce jeune homme soit précisément devenu formateur à l'internet à sa sortie de prison...

Plutôt que de falsifier les textes ou de les tronquer, n'est-il pas possible de sortir une page ou une partie d'un site de l'index des principaux outils de recherche, dont les robots suivent généralement des protocoles d'exclusion standard ?

La suppression peut même être quasiment instantanée.

Cela semble être une solution facile à mettre en oeuvre pour ce genre de demandes mais suppose la bonne volonté des gestionnaires des contenus indexés.

[Exemple - supprimer du contenu de l'index de Google](#)

[Répondre à ce message](#)

5. *5 gag 21 avril 2003 @ 19:38*

> le droit à l'oubli | googleries

Le biographe de Libé, Jean Guisnel raconte bien comment les chroniques judiciaires du journal ont été mises en place fin 70 pour dénoncer la justice de classe, qu'elles ont été lues, appréciées voire copiées (par Le Monde) par la pertinence du regard qu'elles portaient sur, par ex., la brutalité des jugements en comparution immédiate pour des crimes de pauvres.

Mais il n'explique pas l'intérêt de donner les noms des personnes jugées (ni des juges).

Effectivement, si les noms originaux des justiciés sont imprimés et conservés dans les articles justice de Libé, ça me paraît une grave connerie. En fait, ça me semble pervertir complètement la mission que ces chroniques s'étaient données au départ

[Répondre à ce message](#)

6. *6 davduf 22 avril 2003 @ 11:15*

l'anonymisation des comptes rendus de presse

Sur la question de l'intérêt de donner les noms des personnes jugées (et des juges) dans les comptes rendus judiciaires de presse (et même si cela nous écarte un peu de l'aspect casier-judiciaire bis via les moteurs de recherches), il est difficile de trancher.

Plusieurs arguments s'entrechoquent. Donner les noms n'a, en effet, qu'une portée informative très très limitée. Dans bien des cas, les initiales ou le prénom + initiale du nom de famille pourraient suffire, notamment dans les cas de justiciables/quidams anonymes.

Probleme, que fait-on quand le justiciable est un personnage public ? Faut-il écrire J.C. ou Jacques C. ? Ou encore Jacques Chirac ? Jésus Christ ,-) ?

C'est tout le probleme entre liberté d'informer / liberté individuelle / droit à l'oubli.

Rappelons juste que la publicité des débats, pour reprendre le terme juridique, qui remonte (je crois) au XIXe siècle, fut, entre autre, une manière pour le législateur d'ouvrir les tribunaux au public et, d'une certaine manière, de les mettre sous l'oeil/surveillance de celui ci... Seulement, l'informatique change la donne...

[Répondre à ce message](#)

7. 7 davduf 4 mai 2003 @ 22:28

A lire aussi

Via Le Journal Du Net

[L'anonymisation des décisions de justice sur Internet](#)

[Répondre à ce message](#)

8. 8 davduf 1er juin 2003 @ 15:42

> A lire aussi

Le dossier spécial Arrêt Sur Image

[Site de France 5](#)

[Répondre à ce message](#)

9. 9 toctoc 2 mai 2004 @ 16:30

> Et le droit à pouvoir rechercher ?

Bonjour Davduf et merci pour ce débat mieux argumenté et plus intelligemment tenu que sur la plupart des forums.

Avertissement : ce qui suit va sembler provocateur à certains.

La CNIL, dans sa délibération n° 01-057 du 29 novembre 2001 portant recommandation sur la diffusion des données personnelles sur internet par les banques de données de jurisprudence <http://www.cnil.fr/index.php?id=1361>, recommande aux éditeurs de supprimer de leurs bases payantes les adresses des parties et aux webmestres des sites gratuits de supprimer noms et adresses. Donc même Legifrance <http://www.legifrance.gouv.fr> le fait depuis juillet 2002. Le stock, lui, attend, du fait du coût, notamment, évalué à environ 1 million d'euros <http://ameli.senat.fr/public/AmeliG...>

Est-il cependant légitime que la justice, qui est rendue publiquement, et qui diffuse ses copies des de décisions sans anonymisation, doive anonymiser ses décisions dès que celles-ci passent sur Internet ?

On parle d'un droit à l'oubli, quid du droit à rechercher et retrouver ? Et qui a demandé aux associations de victimes, aux syndicats de salariés ou aux associations d'utilisateurs de ces banques de données leur sentiment exact sur la question ? La CNIL, lors des auditions préalables à sa recommandation sur les banques de données de jurisprudence, ne les a pas consultés. Les éditeurs juridiques, en revanche, ont été consultés (cf le visa de la délibération : >> Entendus [...] des représentants des Editions ..., ..., ... <<). Je constate d'ailleurs que cette recommandation ne demande aux bases de données payantes — appartenant aux éditeurs pour 95% d'entre elles — que d'anonymiser les adresses des parties, pas leurs noms.

Les décisions, surtout celles qui font jurisprudence, sont depuis des lustres, désignées par le nom de l'affaire. Or celui-ci n'est, dans 99% des cas, que le nom d'une partie ou des deux parties. C'est de fait le seul moyen à la disposition de la mémoire des juristes. Et en ce qui concerne les publications des éditeurs et leurs bases de données, il serait temps de dire clairement que le taux d'erreur dans les dates et numéros des arrêts cités se situe entre 3 et 5%

en France. En revanche, sur les noms, quasiment pas d'erreur. S'il y en a, leur racine est le plus souvent correcte et on peut ainsi ne chercher que sur elle ou corriger de soi-même. D'où l'intérêt pour les gens du monde juridique de conserver la possibilité de chercher par le nom des décisions, c'est-à-dire celui des parties.

Enfin, quoiqu'on fasse, si l'on est une personne connue un tant soit peu, les 3/4 de nos relations sauront que l'on a fait l'objet d'une décision. Les réclamations des condamnés en justice sont parfois assez hypocrites et ne régleront que 50% de leur problème de mauvaise réputation. De plus le bouche à oreille reste terrible et rien n'empêche l'avocat d'un employeur d'interroger Lamyline, Juris-Data/LexisNexis.fr ou Juripro, bases de données payantes, où les noms restent.

Il serait intéressant d'équilibrer les intérêts de manière à anonymiser ce qui est vraiment dangereux : les décisions pénales (crimes et délits), les divorces, la copropriété et la location immobilière, les décisions sur les mineurs, etc. D'ailleurs, avant l'apparition des bases de données puis d'Internet, les décisions pénales étaient déjà largement anonymisées à coup de X... et Y... Si les décisions de justice étaient au format XML — mais l'informatisation de la Justice a un tel retard et une qualité tellement faible... — une simple balise et le tour serait joué.

On pourrait aussi ajouter à cette anonymisation par matière une anonymisation à la demande du juge ou d'une partie.

Des gens très sérieux, dont un avocat, plusieurs magistrats et même la Commission belge de la protection de la vie privée, ont publiquement pris position pour une telle anonymisation sélective et à la demande. Je rappelle également que les décisions de la Cour européenne des droits de l'Homme (CEDH) ne sont pas anonymisées sauf lorsque cela est demandé par les parties.

Mais sur un sujet qui en France porte à fantasmes depuis l'affaire SAFARI, et d'autant plus que les Français n'y connaissent pas grand' chose en risques réels de l'informatique, je demande peut-être trop...

.....

3 - Quelques services, logiciels ou sites liés à la thématique et qui ont fait l'objet d'une brève sur le site dreamorange.fr

Aout 2007, rédigés par Hélène Caubel

Internet Archive : une gigantesque bibliothèque numérique libre d'accès

Depuis 1996, dans une dynamique non marchande, le site états-unien Internet Archive répertorie, avant tout des pages web, mais aussi des textes, des films, des enregistrements de concert, et plus généralement tout type de documents, pourvu qu'ils soient numérisés. Les archives ainsi constituées se comptent en milliards. Le site s'adresse aux chercheurs, aux historiens, aux étudiants ainsi qu'au grand public. Il se revendique comme outil universel d'accès à la connaissance. Au-delà de sa fonction de bibliothèque numérique, Internet Archive permet d'appréhender la fulgurante évolution du web en termes de contenus, d'ergonomie, de design. Il semblerait que le site soit également utilisé à des fins juridiques afin de prouver l'existence, à une date donnée, de contenus en ligne prêtant à litige. En septembre 2006, sous l'impulsion d'Internet Archive, une [version européenne](#) a vu le jour. Infiniment plus pauvre sur le plan de la documentation proposée - une centaine de pages web sont présentées contre 85 milliards sur Internet Archive - le site paraît cependant mieux structuré.

Dandelife.com, des vies en ligne

Kelly Abbot, fondateur de dandelife.com, aime jouer avec les mots. Le nom qu'il a attribué à son site évoque le pissenlit : *dandelion* en anglais. La tradition populaire veut en effet que le nombre de souffles nécessaires pour que s'envolent toutes les graines de la fleur corresponde à l'heure qu'il est. Dans certaines campagnes, plus ce nombre s'élevait plus la perspective de se marier s'éloignait. Le surnom du pissenlit, à savoir *tell-time*, témoigne très

explicitement de ce rapport au temps. Kelly Abbot revendique également la paternité du mot *Lifecasting*, désignant le fait de diffuser sa vie sur le net en continu à travers du texte, des photographies, des vidéos. Cette pratique est largement antérieure au lancement, en version bêta, de dandelife.com en juillet 2006. Le principe de ce site n'est donc pas révolutionnaire malgré ses innovations langagières revendiquées. Il s'agit d'un réseau social en ligne, reposant sur les récits de vie de ses utilisateurs (tels Ourstory.com ou encore Immortalspace.com). Ces récits sont placés sur une frise chronologique. Ils échappent ainsi à la présentation habituelle du blog, billet par billet. Ils peuvent être agrémentés par un simple « glisser-déposer » de photographies et de vidéos personnelles ou non, en provenance de Flickr et de YouTube. Il est également possible de *tagger* ses propres histoires pour qu'elles puissent faire écho à toutes les autres présentes sur le site, d'agrèger des flux RSS sur sa page, de faire référence à ses autobiographes favoris sur le site... Kelly Abbot déclarait, dès 2006, que ce vaste journal extime (plutôt qu'intime) pourrait servir non seulement de terrain d'étude pour les annonceurs mais aussi de support de publicité. Dans les récits de vie des internautes se glisseraient alors très naturellement des marques : celles de leur robe de mariée, de leurs premières chaussures, de leur voiture ou encore des boissons qu'ils boivent pour se requinquer.

Ziki, la mémoire recomposée et maîtrisée

Ziki.com est un site français, lancé en avril 2006 en langue anglaise, afin de conquérir plus aisément le vaste monde. La version française a vu le jour en novembre de la même année. Sur ce site, il est possible d'agrèger en un seul et même endroit l'ensemble des flux d'information que vous avez produits et *taggés*, de façon éparse, sur le web : billets et commentaires publiés sur des blogs, photographies, vidéos déposées ici ou là, liens del.icio.us, etc. Votre parcours numérique, les gestes et les contenus qui y sont associés, peuvent être appréhendés dans leur globalité... ou en partie. Car il vous est possible de choisir ce qui vous semble le plus opportun de retenir et de présenter. Il est également possible de compléter les informations recensées en ajoutant du contenu directement sur le site. Ziki.com est un outil d'archivage, de centralisation et de mise en valeur de votre activité numérique. Il est d'ailleurs avant tout présenté comme un « gestionnaire d'identité ». Si quelqu'un recherche votre nom sur Internet et que vous vous êtes constitué un profil Ziki, une publicité s'affiche automatiquement signalant : « Vous voulez tout savoir sur moi. Consultez ma page personnelle ici ». Le site revendique aussi d'appartenir à la grande famille des réseaux sociaux sur le net. Il vous offre ainsi de découvrir d'autres membres avec qui vous partagez des affinités. Le tout nouveau Ziki, sorti en juin 2007, vient renforcer ces dimensions « identitaire » et « sociale » en proposant un moteur de recherche spécialisé dans les informations sur la personne, une fonctionnalité permettant de faire des recherches sur le contenu d'un Ziki, un récapitulatif de vos Ziki favoris et de vos fans (c'est-à-dire ceux qui vous ont répertoriés dans leurs propres favoris) ou encore la republication du CV que vous avez déposé sur le puissant réseau social LinkedIn.

Nokia Lifeblog : se fabriquer des souvenirs... mobiles

Associant un logiciel pour PC et un logiciel pour téléphone mobile, Nokia Lifeblog permet de réaliser, de faire évoluer et de partager, sur l'un et l'autre support, son journal personnel fait de textes, de photographies, d'enregistrements sonores, de vidéos. Lorsque le mobile est saturé de souvenirs, il est ainsi possible de les transférer sur PC. Comme il est possible de sélectionner les « morceaux choisis » de son blog sur PC et de les transférer sur mobile. Un classement par ordre chronologique des différents éléments collectés, créés, envoyés ou reçus s'effectue automatiquement. Des informations complémentaires peuvent être associées à ces éléments par voie audio ou textuelle. Cette organisation dans le temps et cette possibilité de marquage spécifique facilitent les recherches ultérieures... lorsqu'il s'agit de se rafraîchir la mémoire. Pour peu que vous emmeniez votre téléphone mobile partout, Nokia Lifeblog favorise la captation sur le vif d'instant de vie, leur narration puis, plus tard, leur convocation où que vous soyez. Ce logiciel est téléchargeable gratuitement mais ne fonctionne qu'avec les terminaux Nokia Nseries.

L'association PING : pour une mémoire numérique participative

L'association PING, créée en 2004 à Nantes, favorise l'accès aux outils et aux pratiques multimédias en mettant à la disposition de structures ou de particuliers - engagés dans des démarches à but non lucratif - des ressources humaines et techniques. Elle conduit son action autour de trois grands axes : l'organisation de séances de formation et d'expérimentation ; l'accompagnement de porteurs de projet (artistes, enseignants, acteurs associatifs...) ; le développement de projets culturels en propre. En 2006, faisant le constat de l'émergence dans le champ du numérique de nombreuses initiatives liées à la mémoire, PING au côté notamment d'acteurs brestois et de l'association de Bug Rennes impulse un « wiki mémoire ». Sur cet espace commun de travail en ligne, chacun est invité à s'interroger sur les enjeux de l'alliance entre mémoire, territoire et numérique et à témoigner de son expérience dans le domaine. En parallèle, l'association épaula l'artiste multimédia Cédric Doutriaux dans la mise en place et la mise en valeur de son installation « Les murmurs du souvenir » (<http://murmurs.ouies.net/>). Le dispositif

restitue des sons, des phrases, des gestes qui y ont été déposés par ceux qui l'ont traversé. Il n'existe qu'une fois « nourri » et « activé » par les visiteurs. Au fil du temps, les traces s'accumulent et resurgissent à chaque nouveau passage. A travers des projets artistiques (NovYPASSé, VideoVox), des rencontres (sur le thème par exemple de l'économie de la mémoire numérique), des ateliers (atelier web-radio, atelier du Sonographe) : PING poursuit avec assiduité sa réflexion sur la collecte, le stockage, la production, la ré-activation ou la fabrication d'une mémoire collective via le multimédia

Lifelog Pod de KDDI, un réseau mobile capable de restituer vos moindres gestes

L'opérateur en télécommunications japonais KDDI a mis au point un serveur capable non seulement de garder en mémoire tout le détail des activités qu'une personne réalise à partir de son téléphone mobile et de son ordinateur mais aussi de leur donner une visibilité sur d'autres téléphones et ordinateurs. Ce Lifelog Pod enregistre et restitue, tout en les localisant et datant, les appels que vous donnez et recevez ainsi que les recherches que vous faites, ou la musique que vous écoutez, les transferts d'argent ou encore les achats que vous effectuez. Votre réseau social (mais aussi le réseau social de votre réseau social) peut avoir accès à l'ensemble de ces informations. KDDI y voit une manière de se faire de nouveaux amis ! Il rappelle, pour ceux qu'une telle perspective de « réseau omniscient » effraierait, que l'utilisateur gère lui-même le droit d'accès à sa mémoire et vie téléphonique et informatique. KDDI a communiqué en 2006 sur cette technologie pour prouver qu'il était en mesure de la mettre à disposition de ses clients mais, à ce jour, l'opérateur n'a pas encore sauté le pas de la commercialisation. Il continue de faire évoluer le système qu'il a développé. Jusqu'ici, celui-ci ne pouvait fonctionner que si l'utilisateur possédait un terminal GPS. L'utilisation du RFID pourrait rendre caduque cette obligation et réduire encore un peu plus le temps de collecte et de restitution des informations.

.....

4 – Un dossier sur la mémoire numérique paru sur le site dreamorange.fr

Mémoire numérique : en route vers l'illimité

Aout 2007, rédigé par Julien Duprat

Les nouveaux territoires de la mémoire

À pratiques numériques... mémoires numériques

Notre mémoire s'apprête à vivre une révolution. Alors que nous passons rapidement du monde de Gutenberg aux requêtes sous Google, la mémoire, elle, s'étend déjà sur de nouveaux supports numériques... Et comme le souligne Michel Serres, philosophe et historien, à chaque fois que l'humanité a su trouver des « supports de mémoire », « des explosions d'invention » ont été observées.

La croissance exponentielle de nos pratiques numériques génèrent un foisonnement de données numériques et créent de nouveaux besoins : comment garder une trace de notre vie numérique et organiser cette mémoire virtuelle ?

La *folksonomy* et le développement des outils de *tagging* ont esquissé une première réponse : nous pouvons aujourd'hui nous créer une prémisses de mémoire numérique sur le Web à travers laquelle nous naviguons par mots-clés. Les services basés sur la localisation, ont permis d'y ajouter une composante spatiale.

Les *lifelogs* constituent la prochaine étape décisive du développement de notre mémoire numérique personnelle. Il s'agit de systèmes qui permettent à un utilisateur de capturer, stocker et rendre accessibles ses flux de données et d'informations issus d'un large éventail de ses expériences et interactions. L'institut de recherche Nomura¹⁰ estime qu'à l'horizon 2010, l'utilisation des *lifelogs* se répandra au sein des utilisateurs grand public.

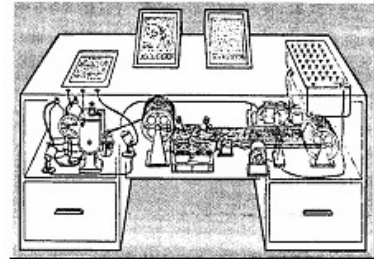
¹⁰ <http://www.nri.co.jp/english/opinion/papers/index.html>

L'extension numérique de notre mémoire, une conception pré-Web

Le concept de *lifelog*, a été exposé pour la première fois à la fin de la seconde guerre mondiale par Vannevar Bush, alors conseiller scientifique du Président Roosevelt et directeur du NDRC (National Defence Research Committee). Ce dernier est aujourd'hui reconnu comme un des pionniers d'Internet, au même titre que Tim Berners Lee ou Vint Cerf.

En 1945, lors de la publication remarquée de son article « As We May Think » et près de 45 ans avant l'invention du World Wide Web, il expose les fondements théoriques du lien hypertexte et pose déjà les bases de Memex¹¹, une extension numérique de notre mémoire. Il décrivait alors une machine reposant sur une technologie de microfilm à partir de laquelle il serait possible de stocker l'ensemble des contenus et informations auquel un individu est confronté. Ce Memex devait être construit dans un bureau, équipé d'un clavier, d'un microphone, d'une caméra frontale de capture des documents et de plusieurs zones d'affichage. Il envisageait alors un système de recherche et de gestion des documents au sein de cet espace de stockage basé sur des « hyperliens » en prenant pour modèle le fonctionnement par association du cerveau humain.

60 ans et une révolution informationnelle plus tard, la construction des mémoires numériques demeure techniquement complexe. En revanche, bon nombre des briques nécessaires à la construction de la vision originelle de Vannevar Bush sont désormais disponibles.



Le Memex de V. Bush (1945)

¹¹ Memex : Memory Extender

Lifelog : vers une mémoire totale

L'abolition des barrières technologiques

Premier élément de poids dans la question de l'extension numérique de notre mémoire : le stockage. La baisse continue du prix du Giga-octet, le développement de nouvelles technologies proposant de meilleures densités de stockage et la multiplication des offres de stockage réseau permettent d'envisager à court terme la possibilité de numériser l'ensemble de sa vie.

La numérisation croissante de nos activités et la forte pénétration d'outils de capture et de communication au sein du grand public (appareil photo numérique, téléphone et ordinateur portable, microphone/casque, webcam) favorisent l'arrivée de telles extensions de mémoire. La démocratisation actuelle des terminaux GPS rend possible la spatialisation de notre mémoire numérique. A long terme, le déploiement de réseaux de capteurs, ainsi que le développement de la biométrie, devraient apporter d'autres sources d'information à nos mémoires numériques.

Enfin, des avancées décisives sont en cours en traitement du signal, multipliant de fait les possibilités d'interprétation et de compréhension de ces futures bases de données personnelles : technologies de recherche audio et vidéo reposant sur des algorithmes bayésiens¹², reconnaissance optique fine des caractères, reconnaissance faciale au sein des photographies, reconnaissance et synthèse vocale.

Ma vie en boîte, ma vie en Bits

A l'instar de Hewlett-Packard ou du MIT qui ont initié des travaux relatifs aux *lifelogs*, une équipe de recherche s'est constituée chez Microsoft autour d'un projet s'inspirant du Memex : MyLifeBits. Véritable base de données historique et personnelle, MyLifeBits requiert l'utilisation d'un système portable de capture, la SenseCam, sorte de boîte noire personnelle programmée pour effectuer jusqu'à 2000 photos par jour. Depuis 2001, Gordon Bell, un des membres de ce groupe projet, s'est constitué son *lifelog*. scanné, capté, numérisé et stocké l'ensemble des informations – photos, livres, pages web, films, sons – auxquels il a eu accès, ainsi que l'ensemble des traces de ses déplacements, de ses échanges et communications.



SenseCam de Microsoft

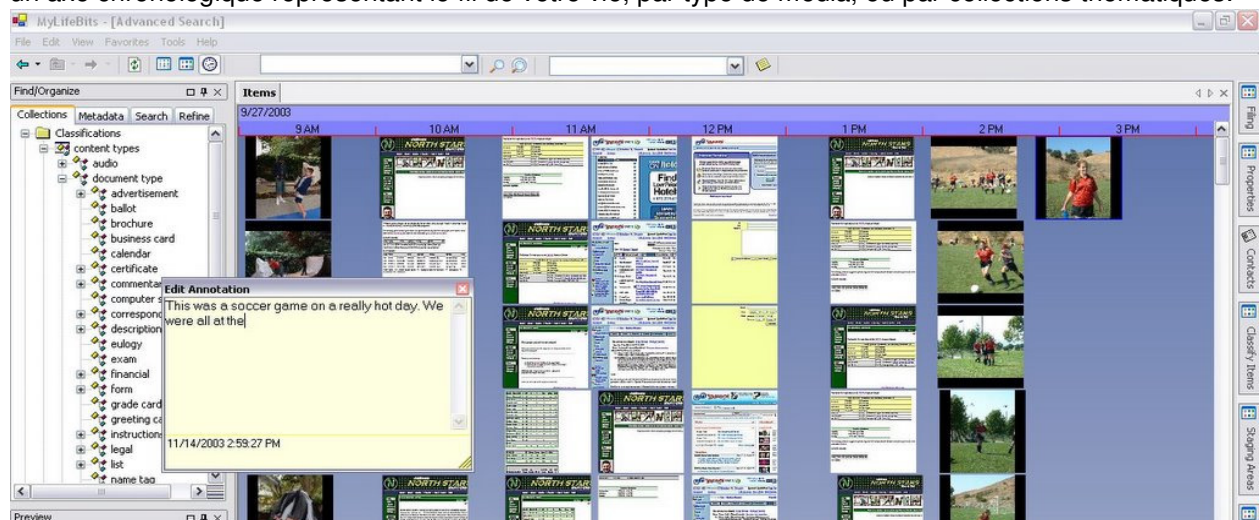
Il a
des

Accumulant près de 2500 Go de données par année, les équipes de Gordon Bell ont vite été confrontées au problème de surcharge informationnelle. Identifiée depuis longtemps par l'industrie informatique, cette surcharge cognitive est en passe de devenir une problématique essentielle dans notre environnement tant professionnel que personnel. Classer, sélectionner et manger ses informations devient un défi majeur.

Afin d'être à même de fournir une interface intelligente capable de stocker toutes nos données personnelles, mais surtout de les retrouver avec rapidité et flexibilité, un des choix technologiques effectué par les équipes de MyLifeBits a été de ne pas créer d'arborescence d'archivage de données.

¹² Les algorithmes bayésiens sont basés sur des méthodes de calcul probabilistes et offrent des estimations très efficaces de la pertinence des résultats lors des recherches. Ces algorithmes sont aujourd'hui fréquemment utilisés dans les logiciels anti-spam.

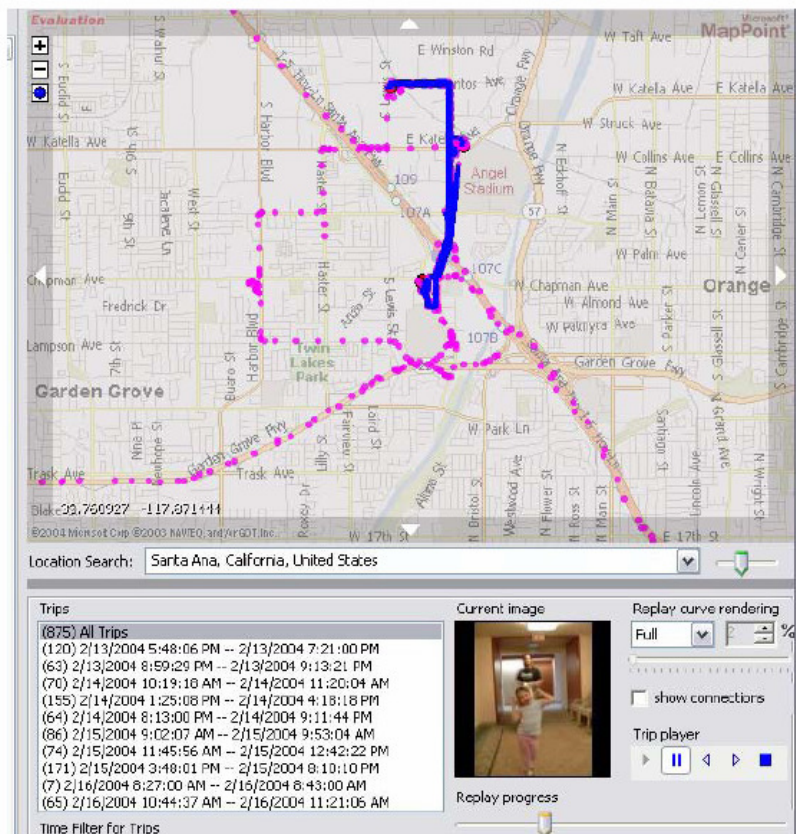
En revanche, plusieurs modes de visualisation des mêmes données sont proposés dans MyLifeBits : sur un axe chronologique représentant le fil de votre vie, par type de média, ou par collections thématiques.



Mais quel serait l'espace de stockage nécessaire pour conserver l'ensemble de nos emails, documents et livres, images, pages web visitées, musiques et sons écoutés, photos provenant de la SenseCam, et contenus vidéos visionnés ?

L'équipe de MyLifeBits estime qu'avec 200 Go, il serait possible de numériser un mois complet de la vie d'une personne, soit près de 200 To pour 83 années.

Au-delà de ces contenus numériques ou numérisables, un capteur personnel permet également de mémoriser des choses dont nous ne sommes même pas conscients, comme les taux d'oxygènes dans le sang, les taux de CO2 dans l'air, le rythme des battements de cœur et autres constantes physiologiques...



La question du stockage résolue, il ne faut pas pour autant omettre celle de l'effacement et donc de l'oubli de certains évènements. En effet, si dans certains cas l'oubli fait suite à un dysfonctionnement, il est généralement considéré comme un « nettoyage mental » indispensable à la santé psychique de l'individu, soit parce que l'information est dépassée ou présente peu de chances d'être à nouveau utile, soit parce qu'elle est traumatisante.

Intimité augmentée...

Promesse d'une mémoire augmentée

Véritable rupture, l'extension numérique de la mémoire humaine constituera probablement dans les années à venir, une source considérable d'innovation en matière de services destinés à faciliter la vie quotidienne. Dès aujourd'hui, de nombreuses applications connexes au concept de *lifelog* sont envisagées.

Ainsi, dans le domaine médical, le *lifelog* pourrait répondre à certains troubles de la mémoire.

Le monitoring des constantes physiques permettra de prévenir certaines maladies ou au moins d'identifier de façon très précoce des modifications du fonctionnement de notre organisme.

L'existence d'une grande quantité d'informations emmagasinée tout au long d'une vie pourrait également bénéficier de l'intégration d'informations provenant d'autres sources de médias afin de créer une narration des événements. Après l'écriture, la peinture, la photographie et dernièrement le cinéma, les *lifelogs* offriront au plus grand nombre la possibilité de transmission d'une histoire personnelle ou collective (mémoire d'une association, d'une entreprise, d'un quartier, d'un parti politique...).

Enfin, grâce à la constitution passive d'une base de données de nos activités, des outils d'*Hyper Life management* pourront voir le jour. Une analyse de l'allocation de nos propres ressources de temps, de nos déplacements quotidiens, des dépenses mensuelles, nous permettra d'améliorer notre organisation personnelle, si nous le souhaitons.

Menace potentielle sur le respect de la vie privée

La multiplication des services en ligne, qui entraîne l'internaute à disséminer des traces numériques toujours plus nombreuses et à exposer certains éléments de sa vie privée, pose avec une acuité renforcée les questions de sécurité et de respect de la vie privée.

Dans ce contexte, la constitution d'une base de données historique et personnelle, ajoute une source de risques supplémentaire.

Les questions de protection et sécurisation de cette base de données, ainsi que la garantie de non-exploitation commerciale devront nécessairement être clarifiées afin de rassurer les utilisateurs potentiels de ce type de service.