



HAL
open science

Vers l'archivage total de soi ?

Jean-Daniel Zeller

► **To cite this version:**

Jean-Daniel Zeller. Vers l'archivage total de soi ?. 12es Journées des Archives de l'Université catholique de Louvain, Apr 2012, Belgique. 20 p. sic_00723875

HAL Id: sic_00723875

https://archivesic.ccsd.cnrs.fr/sic_00723875v1

Submitted on 15 Aug 2012

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.

Actes des 12^{ème} Journées des archives de l'Université Catholique de Louvain (UCL)
De l'individu à la série - Les archives personnelles : enjeux, acquisition, valorisation
Louvain-la-Neuve les 19 et 20 avril 2012, (à paraître en mars 2013)

Vers l'archivage total de soi ?

Une mémoire qui n'opère que dans le passé n'a rien de bien
Lewis Carroll, De l'autre côté du miroir

1. Introduction

Un certain nombre d'évolutions et d'expériences dans l'utilisation des outils informatiques en relation à la personne ont vu le jour ces dernières décennies. Celles qui occupent le devant de la scène médiatique sont celles en rapport avec les réseaux sociaux, ce qui se comprend si l'on admet qu'elles ont une affinité certaine avec le monde de la communication. L'objet de cette présentation est d'examiner un aspect plus discret, quoiqu'également présent sur le web, qui est celui de l'archivage de soi, à travers trois de ses manifestations :

- le Lifelogging (journal de vie)
- le Quantified Self (moi quantifié)
- la Digital legacy (patrimoine numérique)

Le fait que les néologismes anglo-saxons n'ont pas encore trouvés leur équivalent dans l'usage francophone est un indicateur de l'aspect encore novateur de ces pratiques. Après avoir exposé la nature de ces activités, on s'interrogera sur l'impact futur de celles-ci dans la société et sur les archives privées telle que nous les entendons aujourd'hui.

Cet exercice semblera à beaucoup de *l'archivistique-fiction* mais je rappellerais néanmoins que toutes les technologies qui ont durablement impactés l'archivistique sont nées en-dehors de celle-ci et pour des besoins qui n'avait primitivement rien à voir avec la conservation des documents, Comme le parchemin, le papier, la machine à écrire, la photocopieuse et last but not least l'informatique.

2. Le lifelogging

Le terme est une invention de Gordon Bell, chercheur chez Microsoft, et étroitement associé à son livre *Total Recall* où il relate son expérience de lifelogger et à une application développée dans les laboratoires de Microsoft sous le terme de *MyLifeBits* (littéralement : « mes bouts de vie », avec un jeu de mots sur le terme *bits* qui est aussi l'unité de base du code informatique).

2.1. Petite généalogie de *Total Recall* (la remémoration totale)

En 1956, Philip K. Dick, auteur de science fiction, écrit une nouvelle sous le titre de « *We can remember it for you wholesale* » qui sera traduite en français en 1984 sous le titre de « Souvenir à vendre ». La nouvelle, qui n'est pas du meilleur de l'auteur, met en scène un individu cherchant à se faire implanter des faux souvenirs d'une vie sur Mars. Mais comme souvent chez Dick, l'opération s'avère foireuse et le héros retrouve des souvenirs antérieurs comme agent secret sur Mars ce qui déclenche une cascade de conséquences désastreuses. En 1990 cette nouvelle fera l'objet d'une adaptation cinématographique à succès sous le titre « *Total Recall* ». La nouvelle sera réduite à un prétexte pour une avalanche d'effets spéciaux et la mise en valeur des biscottos de M. Schwarzenegger.

Enfin Gordon Bell utilisera le terme *Total Recall* comme titre de son livre, édité en 2009 avec le sous titre *How the e-Memory Revolution will change Everything*¹. Il sera réédité en 2010 sous le titre *Your Life Uploaded: The Digital Way to Better Memory, Health, and Productivity*. Probablement pour éviter quelque procès avec les producteurs du film.

2.2. Qu'est-ce que le lifelogging ?

Le lifelogging a été défini en février 2007 dans un article de Kevin Kelly² comme : ***Enregistrer et archiver toutes les informations de sa vie. Cela comprend tous les textes, toutes les informations visuelles, tous les fichiers audio, toute l'activité médias, ainsi que toutes les données biologiques provenant de capteurs sur le corps. Les informations seraient archivées pour le bénéfice du lifelogger, et partagées avec d'autres à des degrés divers, et contrôlées par lui seul.***

Pour Gordon Bell, son défenseur le plus éminent, l'utilisation d'auxiliaires informatiques pour enregistrer le plus d'aspect de sa vie est un moyen de développer une mémoire plus complète que celle, défaillante, produite par la biologie humaine. Il estime qu'elle est plus objective, plus durable et plus facilement manipulable (recherchable) par les moyens techniques actuels.

Dans son livre, après avoir exposé les moyens mis en œuvre dans son expérience de *MyLifeBits*, il détaille les avantages que procure cette mémoire technologique dans différents domaines de l'activité humaine, le travail, la santé, l'éducation, et la vie post mortem. Dans les derniers chapitres il expose les menus problèmes techniques qui peuvent se poser et la manière d'y pallier avec les moyens actuels. Et enfin il esquisse dans son dernier chapitre sa vision pour l'avenir en vue des nouveautés technologiques qui s'annoncent.

¹ **Total Recall**, Gordon Bell et Jim Gemmel, Flammarion, 2009.

² **Lifelogging, An Inevitability**, blog The Technium, Kevin Kelly, 27 février 2007, http://www.kk.org/thetechnium/archives/2007/02/lifelogging_an.php [consulté le 3.3.2012]

2.3. Florilège du lifelogging selon Gordon Bell

J'ai rassemblé ici quelques extraits significatifs du livre de Gordon Bell, que je commente en italique.

Toutes ces données peuvent être automatiquement enregistrées et sauvegardées, indexées puis rangées dans un dossier. Vous pourrez ensuite les classer selon des critères de temps, de lieu, ou d'autres liens logiques, pour mieux le retrouver plus tard et passer au crible vos comportements et vos habitudes. (Total Recall, p. 31)

Bell exprime ici une forme idéale de la description archivistique, on verra plus loin que cette vision idyllique est loin d'être praticable dans la réalité.

Les gens qui dévoilent ainsi leur vie sur la Toile sont des *life bloggers*. Moi je suis un *lifelogger* : je saisis ma vie entière dans mon e-mémoire. Au risque de paraître vieux jeu, je trouve inconséquent de livrer trop d'information sur soi, surtout à un public aussi large et sans restriction d'accès. (Total Recall, p. 38)

Ce souci de préserver sa vie privée a conduit Bell à négliger l'aspect archivistique de sa démarche, dans une espèce d'illusion d'immortalité qu'une telle démarche implique, nous en reparlerons à propos de l'héritage numérique.

Je suis parfaitement conscient du préjugé nourri par nos sociétés contre cette idée, sorte de huitième péché capital, Tout conserver ferait de vous un être obsédé de son passé ; mieux vaut avancer, débarrasser son grenier de toutes ses vieilleries. Dans un monde analogique, ce type de raisonnement est tout à fait sensé. Mais dans un monde numérique, où les contraintes temporelles et financières s'effondrent une à une, la donne n'est plus la même. Pour tout conserver, on a plus à passer sa vie devant des masses de documents et d'objets à classer. Alors, plutôt que de les jeter, autant les numériser (Total Recall, pp. 46-47)

Bell ici encore fait preuve de techno-optimisme auquel les tenants du « full text » nous ont habitués depuis longtemps. Le refus du choix et de la sélection donne l'illusion que tout peut être à portée de clic. Cependant :: Google y arrive mais pas à tous les coups (le moteur de recherche a été un palliatif aux liens « bêtes » et non qualifiés, contrairement au souhait initial de Tim Berners-Lee)

Quand on grave des données sur un CD, l'opération est en général définitive : on parle de WORM (*Write once, read many* : écrire une fois, lire tant qu'on veut). Jim [Gray] disait que j'avais inventé le WORN, *Write once, read never* : écrire une fois, ne jamais relire. Pour toute réponse, je dus admettre que l'absence d'une dénomination pertinente des fichiers était préjudiciable au projet. Je commençais à saisir l'ampleur du problème : le but du jeu était moins de stocker les données que de réussir à les retrouver. (Total Recall, pp. 53-54)

Qui d'entre-vous a récemment consulté les supports de cours de leur jeunes années d'université ? Pour être honnête environs 2%.

Revenons à 2001 : la question du stockage et de l'organisation des données, j'en avais bien conscience, était loin d'être réglée. Nous venions de remplir **seize gigaoctets** de documents et de photographies avec un système de classement imparfait (dossiers, sous-dossiers, etc.) et nous ne disposions d'aucune méthode efficace pour y effectuer des recherches, trier les données, les annoter, les relier entre elles, ou les analyser pour y repérer des constantes.

(Total Recall, p. 61)

*Bell est clairvoyant quant aux difficultés de l'indexation et de la recherche mais tout au long de son ouvrage il compte systématiquement sur la technologie pour répondre à cette interrogation sans aborder la question essentielle de la nécessité de l'oubli (voir les commentaires à propos de **Delete** de Viktor Mayer-Schönberger au point 6.2).*

Tu fais 16 giga, ça te donne quoi ?

Un jour de plus vers quatre planches de bois

Et pour de consoler de voir c'que tu vois

L'idée que ton père l'a vu avant toi

(d'après « 16 tons », Merle Travis, traduction française de Jacques Bertola)

La métaphore de MylifeBits



Source : Fast company, novembre 2006

<http://paleo-future.blogspot.com/2007/02/most-well-documented-lives-in-history.html>

Dans la réalité de Gordon Bell...



Source: <https://www.viconrevue.com/home.html>

De fait, le seul élément technologiquement nouveau que Bell a employé est la SenseCam. Cette caméra prend une image toutes les 25 seconde ce qui lui permet de constituer une sorte de film avec peu de place mémoire occupée.

2.4. Le lifelogging comme défense de la vie privée...

Je livre ici un petit ex-cursus. Pour la plupart des gens de la génération des baby-boomers, la vie privée veut encore dire quelque chose, et nous sommes plutôt interloqués de voir les nouvelles générations s'en préoccuper comme d'une guigne. En fait, comme le montrent les travaux de Danah Boyd (<http://www.danah.org/papers/>), les nouvelles générations n'ignorent pas les notions de vie privée mais placent sa limite ailleurs. Et parfois ces limites passent à de drôles d'endroits comme le montre la posture de Hasan Elahi. Ce dernier, artiste et professeur à l'université du Maryland mais d'origine bengalie, a été arrêté en 2002 sur le soupçon de détention d'explosifs. Outré de ces soupçons et des questions inquisitoriales qui lui ont été posées alors, il a décidé de communiquer systématiquement sa vie (à destination du FBI) par une forme d'une auto-surveillance systématique.³

Dans ce cas la transparence s'exerce ici et maintenant et n'a pas de prétention à l'archivage, bien que des images de jours anciens soient consultables sur son site, mais s'affichent selon son bon vouloir et ne sont pas recherchables en tant que telles. Elles sont affichées automatiquement par thématique (repas, WC, transport, magasins, etc.) au gré des clics.

³ Son site est consultable à l'adresse suivante : trackintransience.net [consulté le 2 mars 2012]. Des commentaires sont disponible aux adresses suivantes : <http://slate.fr/lien/45759/hasan-elahi-fbi-vie-privee-surveillance> (oct 2010) et <http://www.internetactu.net/2011/022/18/peut-on-devenir-anonyme-en-publiant-tout-de-soi/> (fév 2011) [consulté le 2 mars 2012].

Ce type d'exhibition a aussi pu avoir un raison narcissique comme le site de jennicam.com qui présentait des images d'une jeune américaine prises automatiquement entre 1996 et 2003, avec des commentaires s'apparentant à un journal, bien avant la vogue des blogs qui ont popularisés ce genres de comportement.⁴

La nouvelle métaphore de MyLifeBits



Source : Reuters

I-Pad mon bel I-Pad, dit moi qui est la plus belle...

3. Le Quantified Self

Le Quantified Self (QS pour les habitués et dans la suite du texte) est également décrit par Gordon Bell, principalement dans son chapitre sur la santé. Mais c'est également un mouvement qui à pris son propre essor et qui a développé une organisation internationale⁵. Le mantra de cette communauté pourrait être : Que faites-vous ? Comment le faites-vous ? Qu'avez-vous appris ?

⁴ Accessible partiellement sur Internet Archives à l'adresse : <http://web.archive.org/web/19980124153412/http://www.jennicam.org/index.html>

⁵ Voir le compte-rendu de la première réunion européenne, <http://www.internetactu.net/2011/12/01/quantified-self-13-mettre-linformatique-au-service-du-corps/>

Contrairement au lifelogging qui s'occupe des données que je perçois ou produit de manière extérieure à moi, le QS s'occupe de capter les données qui concernent mon intérieur. Raison pour laquelle il se concentre principalement sur des données biologiques. Le QS reste avant tout une quête de soi qui finalement, cherche rendre l'informatique personnelle encore plus personnelle, puisqu'elle vise à se connecter, non plus au village global, mais à ce que l'on a de supposément le plus intime : son corps.

Dans les exemples rapportés lors du récent congrès européen du QS je relève les expressions suivantes.

L'étude de Kiel Gilleade lui a permis de montrer plusieurs choses intéressantes. D'abord, en terme de résultats, l'impact profond de nos excès de nourriture et d'alcool sur le rythme cardiaque. Les excès de Noël par exemple perturbent durablement le rythme cardiaque de notre sommeil, qui s'élève au niveau de notre activité diurne.

C'est assurément un moyen pour se rappeler qu'on est responsable de soi. "C'est également un moyen de se battre contre toutes ces entreprises qui nous tracent, qui accumulent des données sur nous : désormais, nous aussi nous avons des données sur nous-mêmes !" C'est une manière de reprendre le pouvoir qu'on nous vole. Enfin, c'est peut-être aussi un moyen de se confronter à "l'autre soi", celui avec lequel nous sommes en conversation permanente : celui que l'on veut être et qu'on n'est pas. Un moyen de discuter entre soi et son comportement, entre ce qu'on fait et ne fait pas. Un moyen de se surveiller, d'être conscient de son comportement. Un moyen pour retrouver son identité, sa place dans ce monde d'objets, de ce que les autres disent de moi... Garder trace de ce que je suis est un moyen finalement "pour se rendre plus puissante". (Laurie Frick)

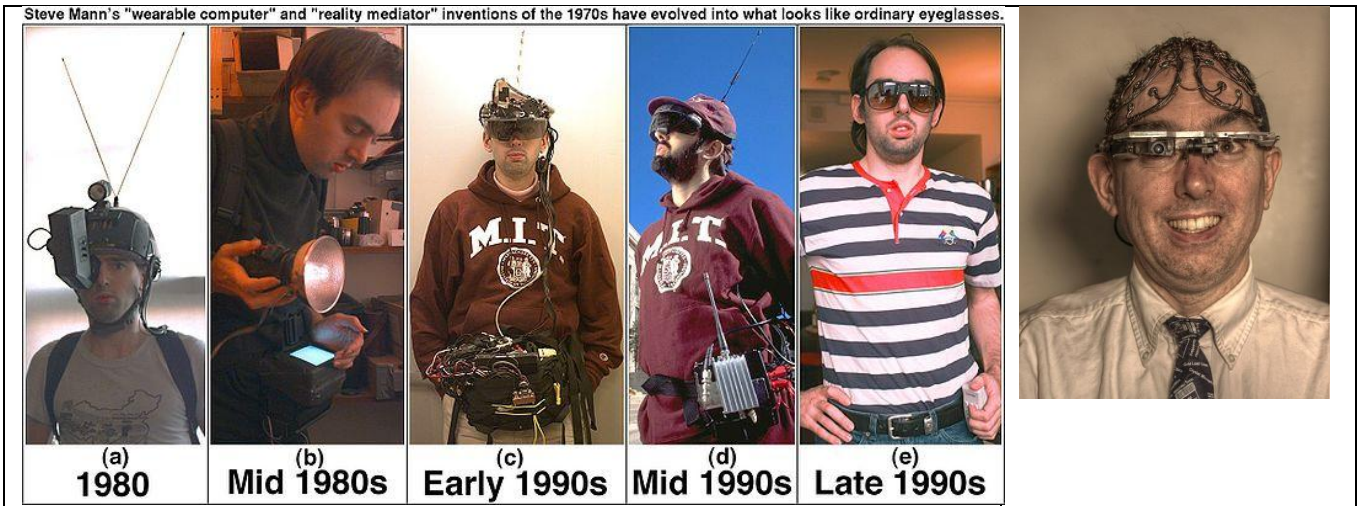
"Quand je vais voir mon docteur et que je lui apporte deux mois de mesure de ma pression sanguine, cela ne l'intéresse pas. Y'a-t-il besoin de reconnaissance des institutions de santé ? Dans le QS on a tendance à construire pour soi, mais peut-être pas en regardant suffisamment au-delà ! Les docteurs ne sont pas payés pour travailler sur les chiffres que je produis. Les modèles d'affaire de la santé sont basés sur des collections de données institutionnalisées, pas sur celles des utilisateurs", regrette le designer (Maarten den Braber).

Ces trois exemples montrent les différentes attitudes que les adeptes du QS peuvent avoir vis-à-vis de leurs données, de la connaissance intime de soi au désir de partage de leur expérience. Pour le QS tout peut être objet de captation et les sites de partage existants rivalisent dans l'exposition des capteurs les plus sophistiqué pour connaître sa vie. Cette dimension de « bricoleur » et d'expérimentateur n'est pas la moindre de ses composante.

L'ultime frontière de ce partage peut être trouvée dans une publication de Adams Hedgecoe et Liz Lyon au 7th IDCC Symposium de Bristol en novembre 2011 sous le titre « Share your

genome ? » [Partager son génome]⁶. Où ils mettent en évidence qu'avec la multiplication des outils d'analyse génétique, ce type de données personnelles, considérées pourtant comme sensibles, risquent d'être diffusées à tout vent.

Un autre aspect problématique est le fait que le QS passe peu à peu d'un équipement assez lourd, demandant une certaine maîtrise technique (voir ci-dessous les images de Steve Mann)...

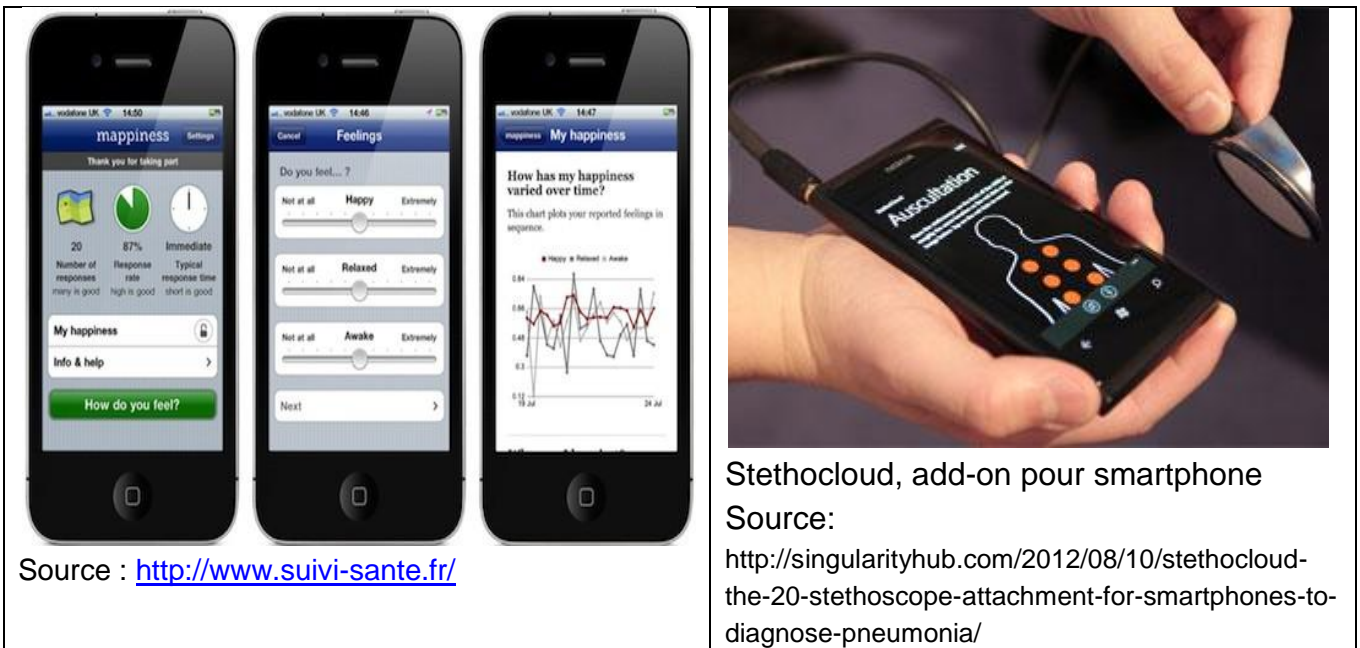


Source :

<http://en.wikipedia.org/wiki/File:Wearcompevolution.jpg>

http://en.wikipedia.org/wiki/File:MannGlassEye_and_MindMesh_2427-2429proc_rotated_cropped.jpg

... à un capteur universel qu'est l'iPhone.



⁶ Share your genome? Dealing with very personal data, Pr. Adam Hedgecoe et Dr. Liz Lyon, 7th IDCC Symposium, Bristol, November 2011, http://www.dcc.ac.uk/webfm_send/608 [consulté le 30.07.2012]

Le smartphone bat en brèche le côté « pirate » et « bricoleur » du QS en mettant à disposition (contre espèces sonnantes et trébuchantes) une forme de captation de soi. Son omniprésence et donc sa facilité d'emploi en tout temps fait oublier que nos données personnelles transitent alors dans le réseau.

4. La Digital Legacy

« Tu es mort, tes données ne le sont pas : qu'arrive-t-il maintenant ? »

C'est le titre d'une réunion qui a eu lieu le 12 mars 2012 à Austin (USA)⁷. Là encore, cette préoccupation de la conservation des données post mortem est plus largement répandue en Amérique du Nord qu'en Europe. On dénombre d'innombrables sites qui, en plus de donner quelques conseils de base sur les mesures à prendre pour prendre soins de ses données en cas de décès, proposent également des services.

Dans ce domaine, les pratiques archivistiques ne sont pas tellement différentes que celles du monde analogique. Simplement, les compétences nécessaires pour gérer l'héritage numérique ne sont pas celles des notaires actuels, comme le remarque ironiquement Pierrot Péladeau⁸. Devant cette méconnaissance, il propose une attitude pragmatique qui consiste à adopter le plan d'action suivant :

Il nous faut donc développer de nouvelles pratiques:

- *les individus doivent apprendre à tenir inventaire de leurs actifs et identités numériques ainsi qu'à signaler leurs volontés à nos proches et nos collègues;*
- *les fournisseurs de services en ligne doivent se doter de politiques sur la disposition des comptes et des contenus par les usagers, leurs mandataires et leur succession;*
- *les entreprises funéraires, les notaires et d'autres doivent offrir un service de référence vers des spécialistes de confiance aptes à aider les exécuteurs testamentaires (ex.: connaissances de la plateforme WordPress, des configurations de médias sociaux).*

Il faut éduquer:

- *les notaires et avocats ; comptables et conseillers financiers doivent se former et s'outiller pour à leur tour éduquer, outiller et soutenir les citoyens de notre monde numérique ;*
- *les banques, compagnies d'assurances, fournisseurs de services en ligne, employeurs doivent aussi nous rappeler la nécessité de gérer ses comptes et contenus numériques.*

Il faut clarifier les dimensions légales:

⁷ **You are dead, Your Data Isn't: What Happens Now?** http://schedule.sxsw.com/2011/events/event_IAP6048 [consulté le 30.07.2012]. Le podcast de cette conférence est actuellement accessible.

⁸ <http://pierrot-peladeau.net/fr/archives/3361> [consulté le 30.07.2012]

- *Des organismes comme les offices de protection des consommateurs, les commissions de protection des renseignements personnels et les ordres professionnels doivent publier des avis et des conseils pratiques pour clarifier l'application des lois et règles déjà existantes. Cela, à l'égard des droits des vivants que de ceux en état d'incapacité ou décédés;*
- *Éventuellement, les législateurs devront réformer certains éléments du droit actuel. Un exemple est ce projet de réforme des directives européennes en matière de protection des informations personnelles afin d'affirmer la capacité des citoyens de retirer ou détruire eux-mêmes des informations non nécessaires de leurs profils sur divers sites.*

D'autres blogs, anglo-saxons encore une fois, développent également des réflexions de ce genre, le meilleur étant certainement *Digital Beyond*.⁹

Les professionnels les plus actifs à propos de cette problématique sont les animateurs du blog de la Librairie du Congrès *The Signal*, blog qui se consacre également, depuis juin 2011, à la problématique plus vaste du *Personal Digital Archiving*¹⁰. La Librairie du Congrès a déjà organisé deux congrès sur ce thème en 2011 et 2012¹¹.

Pour mettre en perspective ce phénomène on peut citer les travaux du philosophe Patrick Stokes qui s'est interrogé sur les similitudes des devoirs envers les morts qui ont de tout temps été virtualisés par toutes les sociétés¹². La pratique de la « mémorialisation » des profils Facebook des personnes décédées, où les gens continuent de dialoguer « avec » ou « autour » des artefacts des gens décédés participent des mêmes modalités. Factuellement 19'000 utilisateurs de Facebook meurent chaque jour. Au rythme actuel, il faudrait 60 ans pour que la moitié des profils de Facebook soient ceux de personnes décédées (sur la base du calcul suivant : 845 millions de profils divisé par 7 millions de décédés par an dans le monde = 120 ans, soit une demi-vie de 60 ans). Il y aura donc de plus en plus de morts-vivants sur Facebook !

5. Les implications sociologiques

Je ne suis pas sociologue, aussi les remarques qui suivent seront celles d'un citoyen moyen s'interrogeant sur l'acceptabilité et les diffusions possibles de ces pratiques, en faisant pour l'instant abstraction de leurs conséquences archivistiques que j'exposerais après. La plupart des éléments que j'expose ici posent plus de questions qu'ils n'apportent de réponses.

⁹ <http://www.thedigitalbeyond.com/online-services-list/>

¹⁰ <http://blogs.loc.gov/digitalpreservation/category/personal-archiving/>

¹¹ <http://www.personalarchiving.com/2012/01/preliminary-schedule-pda-2012/>

¹² **Ghosts in the Machine: Do the Dead Live on in Facebook?** Patrick Stokes, Philosophy and Technology 26.10.2011 <http://www.springerlink.com/content/v258545u7v4h6407/fulltext.pdf> [consulté le 30.07.2012]

J'emprunte un état des lieux sur la question à l'excellente présentation d'Olivier Ertzscheid, qui a abordé cette question lors d'une présentation à l'INA en novembre 2011¹³. Il propose des pistes de réflexion quant à la sémantique qui permet de clarifier quels sont les ressorts qui gouvernent les divers avoires numériques :

- Google, Apple et Amazon STOCKENT les livres que j'ai lus.
- Archive.org CONSERVE mon vieux sky blog d'il y a dix ans
- Facebook DISPOSE de la trace des mes conversations d'il y a 8 ans.
- La bibliothèque du Congrès ARCHIVE Twitter

Plus loin, il propose des points de vue (sociologiques) pour répondre aux différentes représentations de la mémoire numérique :

- Poser la question de LA mémoire DU web a un sens. Le point de vue est PATRIMONIAL.
- Poser la question de LA mémoire DE Facegoog en a un autre. Celui de la privauté des espaces semi-publics. Le point de vue est celui de la CULTURE NUMÉRIQUE.
- Poser la question de NOS mémoires SUR Facegoog. Celui du droit à l'oubli. Existe depuis le projet Safari des années 70 (en France). Le point de vue utilisé est (devrait être ?) LÉGISLATIF.
- Poser la question de NOS mémoires sur LE web. La légitimité du droit de ne pas se souvenir. La question de la transparence. Le point de vue est POLITIQUE.

On perçoit ici une problématique sous-jacente qui fait qu'un même type d'avoir numérique ne devrait pas être traité de la même manière selon qu'il est collectif ou individuel. Ertzscheid l'illustre par une dialectique particulièrement parlante : Julian Assange vous livre gratuitement des informations privées d'entreprises et c'est un affreux ; Marck Zukerberg livre vos informations privées pour de l'argent à des entreprises et c'est l'homme de l'année.

Au-delà de la problématique de la vie privée sur Internet on peut aussi se poser la question de savoir si tout un chacun veut passer sa vie à l'enregistrer comme l'a fait Gordon Bell ou comme le font les usagers du Quantified Self. Mon pronostic est que tant que ces outils restent compliqués à manipuler, cela restera une « geekerie » et ne devrait pas avoir d'incidence majeur, mais on la vu, les fonctionnalités attachées aux smartphones pourraient bien diffuser certain aspects du QS bien plus vite.

5.1. « Devoir de mémoire » versus « Droit à l'oubli »

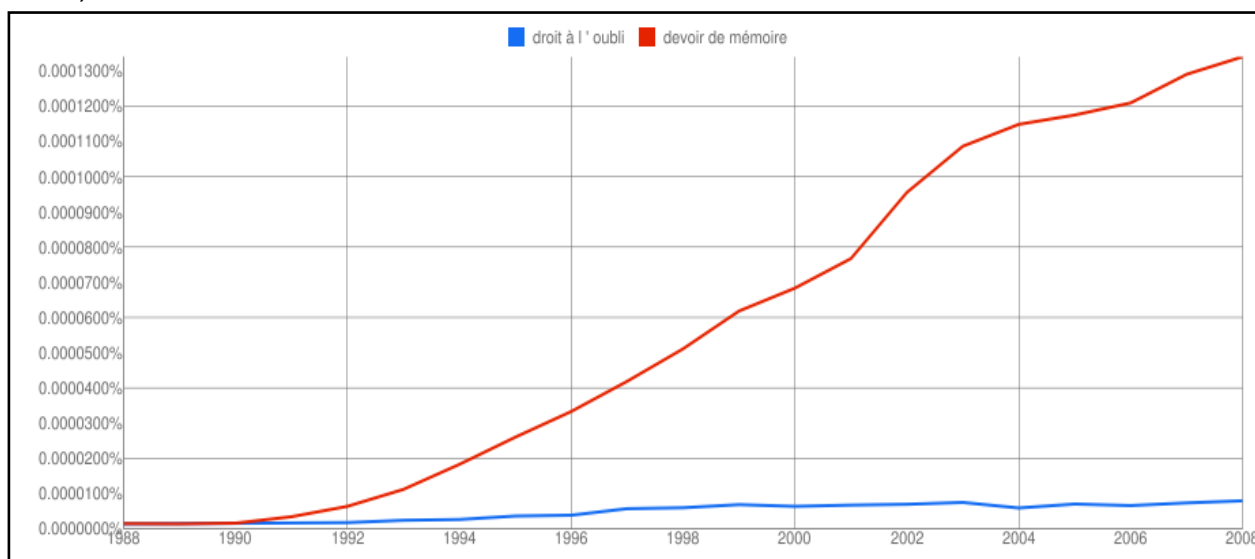
¹³ INA, Ateliers – Dépôt légal du Web, 25 novembre 2011, Olivier Ertzscheid, *Web, grand nombre et archive « la mémoire et l'amer »*, <http://www.slideshare.net/olivier/memoireinaatelierdepotlegalweb>

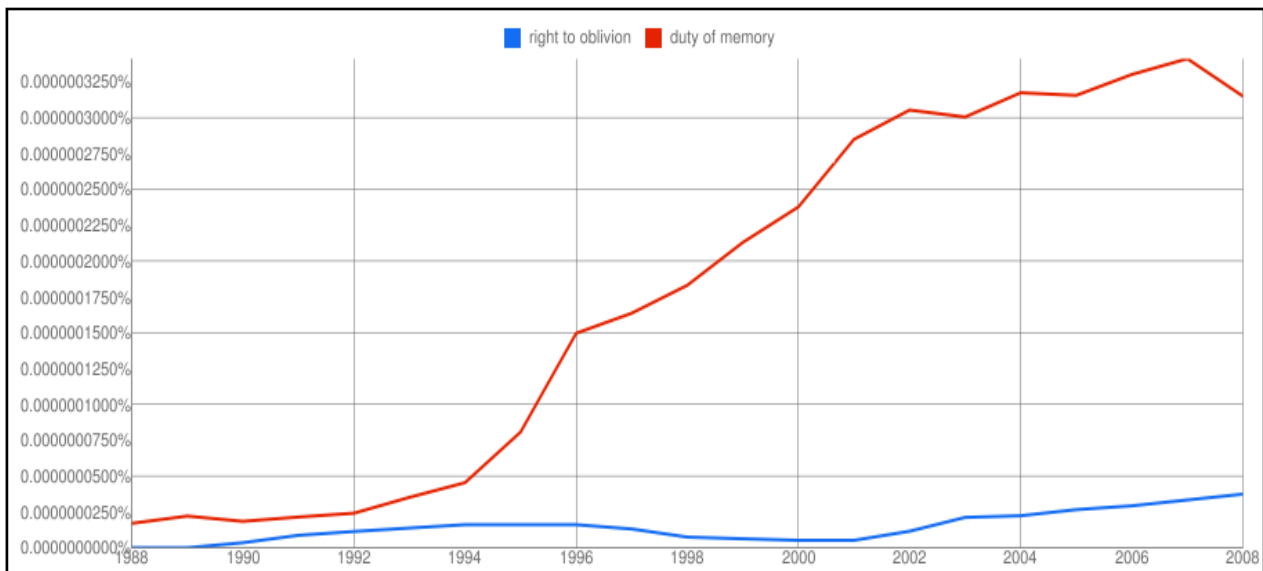
Pour tenter de prévoir si cette mémorisation intégrale a une chance de s'imposer, j'ai modestement tenté d'interroger les oracles googeliens. Les deux expressions « devoir de mémoire » et « droit à l'oubli » nous sont familières et j'ai tenté d'en mesurer le poids réciproque. Les résultats sont les suivants (interrogation effectuée le 16.04.2012). Pour examiner s'il y avait des biais liés aux niveaux intellectuels j'ai procédé à la même interrogation sur Google, Google Livres et Google Scholar. Une interrogation avec les termes en anglais indique une présence beaucoup plus faible des occurrences mais une même tendance en ce qui concerne les proportions.

Français		OCCURRENCES			
Expression	Google	Livres	Scholar	moyenne	
Droit à l'oubli	3'940'000	30'900	410		
Devoir de mémoire	17'200'000	416'000	4'430		
Ratio	4.4	13.5	10.8	9.5	

Anglais		OCCURRENCES			
Expression	Google	Livres	Scholar	moyenne	
Right to oblivion	83'600	75	68		
Duty of memory	161'000	376	344		
Ratio	1.9	5.0	5.1	4.0	

Pour en saisir l'évolution dans le temps j'ai interrogé le Ngram viewer (analyse de Google Book) avec les résultats suivants :





On voit que tant dans le monde francophone qu'anglo-saxon le « devoir de mémoire » a pris son envol dans les publications à partir des années 90. Cette tendance est restée constante et le rapport de 1 à 9 (monde francophone) et de 1 à 4 (monde anglophone) montre une prééminence forte de la mémoire contre l'oubli. Le fait que l'on écrive plus de signifie pas forcément que la valeur de ces expressions est plus partagée sociologiquement, mais il serait intéressant d'explorer plus finement ces résultats. Là également, il faut nuancer entre le domaine privé et le domaine collectif. Le devoir de mémoire s'adresse plutôt à l'obligation collective de se souvenir, alors que le droit à l'oubli fait plutôt en référence au domaine privé.

6. Les implications archivistiques

6.1. Le défi volumétrique

Sachant qu'un individu moyen a déposé environ 1,3 Méga-octets de contenu sur Facebook et que ce dernier compte environs 500 millions d'utilisateurs (en 2010) cela nous donne 650 Téra-octets de données personnelles disponibles sur Facebook (65 fois la bibliothèque du Congrès)¹⁴

650 Téra, et moi et moi et moi...

Avec ma vie, mon petit chez moi

Mon mal de tête, mon mal de foie

J'y pense et puis j'oublie,

C'est la vie, c'est la vie.

(d'après 700 millions de chinois, Jacques Dutronc et Claude Lanzmann)

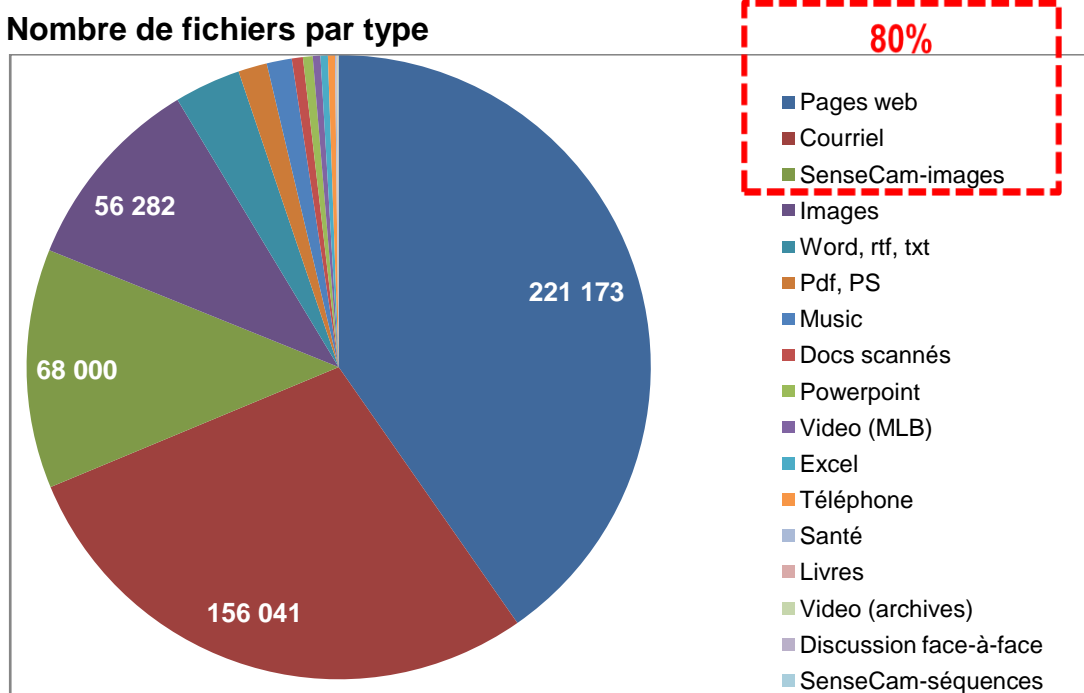
¹⁴ Source : **650 téraoctets de données personnelles, et moi et moi et moi ?**, Olivier Ertzscheld, Affordance.info, 14 novembre 2010, http://affordance.typepad.com/mon_weblog/2010/11/recuperation-donnees-personnelles-facebook.html

Une caméra frontale enregistrant notre vie pendant 70 ans représente **27,5 To** par personne, soit 450 iPod de 60 Go. (source : voir note 13).

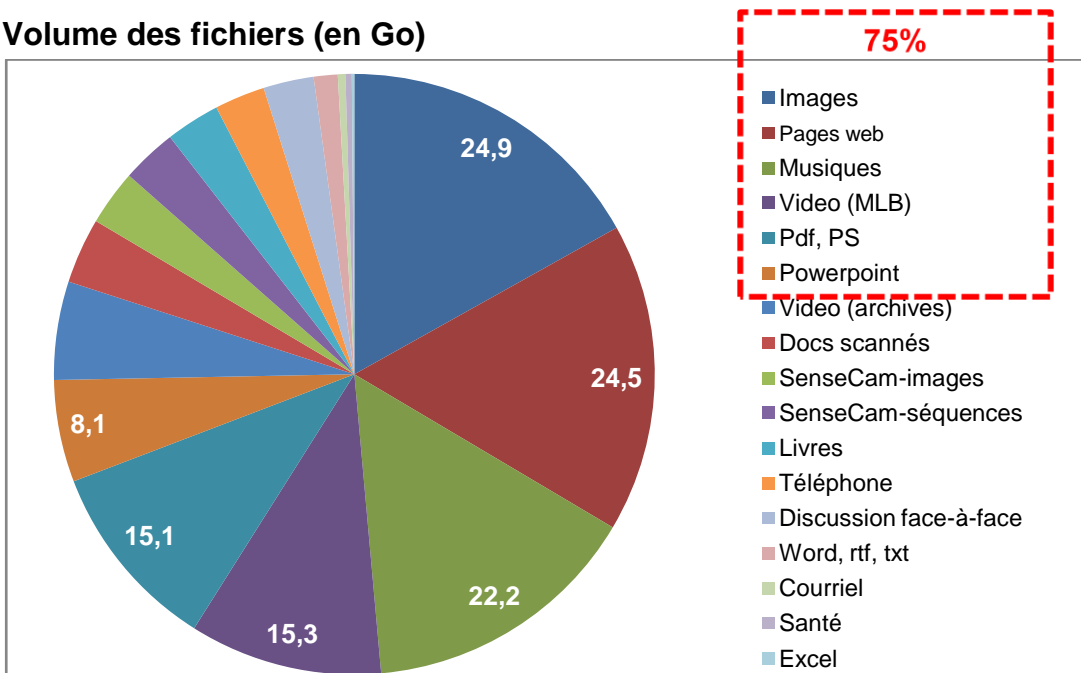
Pour *MyLifeBits*, il s'agit de 580'000 fichiers représentant environ 150 Gigaoctets, dont voici le détail en termes de type de fichiers.

Quel est le contenu d'une vie de bits ?

Nombre de fichiers par type



Volume des fichiers (en Go)



La métaphore finale du lifelogging ?



Magritte, La reproduction interdite, 1937

Si de plus en plus d'humain sont en possession d'avoir numérique les plus divers, que cela soit sur des serveurs commerciaux ou sur leur machine propre, nous atteindrons fatalement des limites. Je n'en cite ici que deux.

Une étude d'IDC¹⁵ estime que vers 2015, le volume de données produites et collectées dépassera la capacité de stockage des supports de données produits par l'industrie (le ralentissement de la production mondiale de ces composants induits par la catastrophe de Fukushima, ne fait que renforcer cette hypothèse).

Une autre étude montre que les « fermes de serveurs » utilisées par les grands acteurs du réseau mondial (Google, Facebook, etc.) consomment une part croissante de l'électricité mondiale (env. de 1% à 2% entre 2000 à 2010, aux Etats-Unis) et que cette tendance va continuer ces prochaines années quoiqu'elle semble se ralentir¹⁶, que se passera-t-il lorsqu'elle dépassera les 10% ?

6.2. Le défi technologique

Dans son article « *The Noise in the Archive: Oblivion in the Age of Total Recall* »¹⁷ [Le bruit dans les archives : l'oubli à l'ère du Total Recall] Jean-François Blanchette porte une critique fondamentale sur le techno-optimisme de Gordon Bell. Sa critique est d'autant plus intéressante qu'il signale le même « déni de matérialité » chez un opposant de Gordon Bell, Viktor Mayer-Schönberger, qui dans son ouvrage *Delete* (non traduit en français)¹⁸, revendique fermement le droit à l'oubli pour les données personnelles. Je cite Blanchette dans sa conclusion :

« Je soutiens que, en dépit de leurs conclusions opposées, les deux auteurs partagent une même vision intellectuelle quant au statut historique unique de l'information numérique comme immatérielle, et donc, imperméable au bruit, à la dégradation et à la distorsion que les supports analogiques apportaient à la tâche de conserver la mémoire. Je conclus que quel que puisse être l'avenir du passé, l'information numérique ne peut pas, en fait, échapper à ses fondements matériels, et va inévitablement amener sur la table sa propre gamme de possibilités et de contraintes particulières à l'entreprise du souvenir et de l'oubli. »

Ce défi technologique n'est pas propre aux données personnelles, mais concerne toutes les données sous forme numérique. Mais la nouvelle organisation « en nuage » de ces données et le discours idéologique qui l'accompagne tend à systématiquement occulter cette

¹⁵ **The Diverse and Exploding Digital Universe** - An Updated Forecast of Worldwide Information Growth Through 2011, John F. Gantz and al., IDC White Paper, March 2008, <http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf> [consulté le 30.07.2012]

¹⁶ **GROWTH IN DATA CENTER ELECTRICITY USE 2005 TO 2010**, Jonathan G. Koomey, Ph.D., Consulting Professor, Stanford University, 01.08.2011, <http://www.analyticspress.com/datacenters.html> [consulté le 30.07.2012]

¹⁷ **The Noise in the Archive: Oblivion in the Age of Total Recall**, Jean-François Blanchette, COMPUTERS, PRIVACY AND DATA PROTECTION: AN ELEMENT OF CHOICE 2011, Part 1, 25-38, DOI 10.1007/978-94-007-0641-5_2, <http://www.springerlink.com/content/p376tpq257816571/abstract/?target=print> [consulté le 30.07.2012]. Traduction française à venir sur le blog "Le regard de Janus" (<http://regarddejanus.wordpress.com/>)

¹⁸ **Delete – The Virtue of Forgetting in the Digital Age**, Viktor Mayer-Schönberger, Princeton University Press, 2009.

matérialité bien réelle, et les risques technologiques qui lui sont associés et que les archivistes connaissent bien.

6.3. Le défi indexatoire

Dans la bibliothèque de Babel imaginée par Jorge Luis Borges, qui contient tous les livres aléatoirement possibles, se pose la question du catalogue de la bibliothèque. Les commentateurs ont argué que, puisque le catalogue devait être alphabétique et que la bibliothèque contient tous les livres, le catalogue ne peut être que la bibliothèque elle-même.

En extrapolant sur cette base le projet de Bell, cela signifierait que pour *MyLifeBits*, il faudrait avoir une nouvelle vie pour pouvoir la consulter dans son intégralité. Gordon Bell ne se cache pas ce problème indexatoire mais il l'évade par la technologie à venir (et les outils de Google lui donnent partiellement raison) et par des outils qui permettront une vision accélérée et/ou synthétique, permise justement par la nature technologique des données.

Nonobstant les béquilles technologiques qui pourraient nous aider à réviser notre vie ainsi enregistrée, je partage l'avis exprimé par l'historien Jean-Pierre Rioux lorsqu'il dit : « Dans un monde où tout fait mémoire, il n'y a plus de mémoire ». Il argumente ainsi : « Nous sommes désormais dans un perpétuel présent où tout est à la fois créé et conservé. Il n'y a plus tri, hiérarchisation et cela rend les historiens très perplexes. S'il n'y a pas de sélection, on arrive à un monde où tout fait mémoire. Autrement dit, à un monde où il n'y a plus de mémoire. Dans ce contexte, le patrimoine n'est plus cet échantillonnage de biens supposés précieux parce qu'ils viennent des pères, mais un flux permanent »¹⁹.

7. Conclusions

- La mémorisation totale (institutionnelle ou personnelle) s'oppose de plein fouet à la notion d'évaluation archivistique.
- Mais elle renvoie aussi au débat actuel sur les *Big Data* qui met en cause le fonctionnement même de la pensée scientifique comme construction d'une théorie à partir de données choisies, pour la remplacer par le traitement algorithmiques de très grandes masses de données.
- La mise en évidence des problèmes pratiques de la captation et de l'indexation des données contient une mise en cause radicale de l'hypothèse précédente (big data).

¹⁹ **Le numérique est-il une menace pour la mémoire individuelle et collective**, interview de Jean-Pierre Rioux, La Croix, 10.01.2012, http://www.la-croix.com/Actualite/S-informer/Sciences/Le-numerique-est-il-une-menace-pour-la-memoire-individuelle-et-collective- NP_-2012-01-02-753137

- Si nous ne pouvons pas compter sur les données (à cause des biais liés à leur représentation) leur capacité de nous livrer des possibilités d'interprétation fonctionnelles est intrinsèquement contestable.
- La mémoire absolue (et/ou la remémoration intégrale) met en cause la pensée historique dans la mesure où celle-ci est une mise en perspective visant à une compréhension synthétique (un modèle).
- Il faut bien entendu distinguer entre une mémoire absolue personnelle et institutionnelle, mais ce n'est qu'une question d'échelle.
- Nous ne serons probablement pas envahis par une pléthore de MyLifeBits.
- Mais certains « geek » célèbres produiront certainement ce genre d'artefact qui seront certainement conservés par des institutions à vocation archivistiques.
(voir Salman Rushdie: <http://elegendre.wordpress.com/2010/03/17/article-fending-off-digital-decay-bit-by-bit/#more-2917>)
- Nous devons certainement développer de nouveaux modes d'évaluation.

A défaut de résoudre ces contradictions, il ne nous restera peut-être plus qu'à chanter cette complainte :

*Que sont mes données devenues
Que j'avais de si près tenues
Et indexées
Elles ont été trop clairsemées
Je crois le temps les a ôtées
L réseau est mort.
Ce sont fichiers que temps emporte
Et que parfois Google rapporte
Et parfois pas.*

*Pauvre sens et pauvre mémoire
M'a Dieu donnée le roi de gloire
Et pauvre archives
Et toutes mes paroles vives
Le temps me vient
Le temps me manque
L réseau est mort
Ce sont fichiers que temps emporte
Et que parfois Google rapporte
Et parfois pas.*
(sur l'air de Pauvre Rutboeuf)

8. Bibliographie

Bell Gordon et Gemmel Jim, **Total Recall**. Flammarion, 2009
http://www.lepoint.fr/science/total-recall-que-la-memoire-vive-11-01-2011-1284646_25.php

Blanchette Jean-François, **The Noise in the Archives : Oblivion in the Age of Total Recall**, COMPUTERS, PRIVACY AND DATA PROTECTION: AN ELEMENT OF CHOICE 2011, Part 1, 25-38, DOI 10.1007/978-94-007-0641-5_2
<http://www.springerlink.com/content/p376tpq257816571/abstract/?target=print>

Ertzscheid Olivier, **650 téraoctets de données personnelles, et moi et moi et moi ?**, Affordance.info, 14 novembre 2010,
http://affordance.typepad.com/mon_weblog/2010/11/recuperation-donnees-personnelles-facebook.html

Ertzscheid Olivier, **L'homme est un document comme un autre**,
<http://archivesic.ccsd.cnrs.fr/docs/00/37/74/57/PDF/articledefOE.pdf>

Gantz John F. et al., **The Diverse and Exploding Digital Universe - An Updated Forecast of Worldwide Information Growth Through 2011**, , IDC White Paper, March 2008,
<http://www.emc.com/collateral/analyst-reports/diverse-exploding-digital-universe.pdf>

Hedgecoe Adam et Lyon Liz, **Share your genome? Dealing with very personal data**, 7th IDCC Symposium, Bristol, November 2011, http://www.dcc.ac.uk/webfm_send/608

Kelly Kevin, **Lifelogging, An Inevitability**, blog The Technium, , 27 février 2007,
http://www.kk.org/thetechnium/archives/2007/02/lifelogging_an.php

Leckart Steven, **Microsoft Researcher Records His Life in Data**, , **WIRED MAGAZINE**: 17.09, 24.08.2009, http://www.wired.com/culture/culturereviews/magazine/17-09/pl_print

Mayer-Schönberger Viktor, **Delete**, the Virtue of Forgetting in Digital Age, Princeton University Press, 2009

Stokes Peter, **Ghosts in the Machine: Do the Dead Live on in Facebook?**, Technologie & Philosophy, 26 September 2011, Springer, DOI 10.1007/s13347-011-0050-7 ,
http://deakin.academia.edu/PatrickStokes/Papers/991983/Ghosts_in_the_Machine_Do_the_Dead_Live_On_in_Facebook

Sites web cités

A propos du Lifelogging

Le site qui vend la SenseCam utilisée par Gordon Bell:
<https://www.viconrevue.com/home.html>

Le site du projet MyLifeBits chez Microsoft : <http://research.microsoft.com/en-us/projects/mylifebits/default.aspx>

Le site de Hasan Elahi, qui documente sa vie heure par heure à l'intention du FBI est là : <http://trackingtransience.net/>, avec un commentaire intéressant ici : <http://www.internetactu.net/2011/022/18/peut-on-devenir-anonyme-en-publiant-tout-de-soi/>

Les archives partielles du site JeniCam (par Internet-Archives), qui rassemble 5 ans de la vie d'une jeune américaine avant la vogue des blogs et des video-blogs est là : <http://web.archive.org/web/19980124153412/http://www.jennicam.org/index.html>

Les billets d'Olivier Ertzscheid sur Affordance.Info : <http://affordance.typepad.com/>

A propos du Quantified Self (soi quantifié) :

Le site international du Quantified Self ; <http://quantifiedself.com/>

Le site du dernier congrès QS européen : <http://quantifiedself.com/conference/Amsterdam-2011/>

Une série d'article reflétant ce même congrès : <http://www.internetactu.net/2011/12/01/quantified-self-13-mettre-linformatique-au-service-du-corps/>

A propos de la Digital Legacy (succession numérique) :

Le dernier congrès américain (mars 2012) Le podcast de cette conférence est actuellement accessible : **You are dead, Your Data Isn't: What Happens Now?** http://schedule.sxsw.com/2011/events/event_IAP6048

Les remarques de Pierre Peladeau : <http://pierrot-peladeau.net/fr/archives/3361>

Les sites qui parlent de cette problématique : <http://www.thedigitalbeyond.com/online-services-list/>
<http://blogs.loc.gov/digitalpreservation/category/personal-archiving/>
<http://www.personalarchiving.com/2012/01/preliminary-schedule-pda-2012/>

Sur le défi volumétrique :

La volumétrie du projet *MyLifeBits*, dans un article paru dans la revue *Wired* (no 17, août 2009) : http://www.wired.com/culture/culturereviews/magazine/17-09/pl_print

INA, Ateliers – Dépôt légal du Web, 25 novembre 2011, Olivier Ertzscheid, *Web, grand nombre et archive « la mémoire et l'amer »*, <http://www.slideshare.net/olivier/memoireinaatelierdepotlegalweb>

GROWTH IN DATA CENTER ELECTRICITY USE 2005 TO 2010, Jonathan G. Koomey, Ph.D., Consulting Professor, Stanford University, 01.08.2011, <http://www.analyticspress.com/datacenters.html>

Sur le défi indexatoire :

Le numérique est-il une menace pour la mémoire individuelle et collective, interview de Jean-Pierre Rioux, La Croix, 10.01.2012, <http://www.la-croix.com/Actualite/S-informer/Sciences/Le-numerique-est-il-une-menace-pour-la-memoire-individuelle-et-collective- NP -2012-01-02-753137>

Sur les conclusions :

La gestion des archives littéraire de Salman Rushdie :
<http://elegendre.wordpress.com/2010/03/17/article-fending-off-digital-decay-bit-by-bit/#more-2917>)