

# TWITTER: UN PATRIMOINE SUPERFLU(X) ?

LE 12 MAI 2010 OLIVIER ERTZSCHEID

La bibliothèque du Congrès (LoC), a récemment annoncé qu'elle allait archiver le réseau de microblogging dans son intégralité, c'est-à-dire même le compte des ihatfuckmondays. Quelles sont les raisons de cette décision ? Quelles questions juridiques soulèvent-elles ?

Ce billet – un peu long – vise à étudier une actualité, l'annonce de l'archivage de l'intégralité du service Twitter par la bibliothèque du Congrès (LoC), et à poser un questionnement sur ce que la nature même de ce type d'archivage change dans notre rapport à une mémoire "collective" et peut être même à la mémoire ... tout court. Le lecteur est averti que les questions soulevées sont beaucoup plus nombreuses que les réponses apportées. C'est parti :-)

**1537.** Par un édit en date du 28 décembre 1537, François 1er "invente" le dépôt légal, qui permettra – beaucoup plus tard tout de même – d'assurer un contrôle bibliographique universel.

**1996. Les pionniers.** Une fondation américaine, l'**Internet Archive** se lance la première dans une tâche d'allure sisyphéenne, l'archivage d'Internet. Un archivage accessible via la machine à remonter le temps du site, la "**wayback machine**". Des milliards de pages web et des millions de documents multi-supports y sont accessibles.

**2001. L'éveil.** Dans un texte de 2001 lors d'un colloque à la BPI\*\*, alors que l'on évoquait le rôle des bibliothèques à l'heure d'une numérisation dépassant à peine le stade de l'artisanat, on posa l'expression et la problématique de la conservation d'un "patrimoine du temporaire". Expression à bien y regarder pas si oxymorique que cela.

\*\*Bibliothèque publique d'information (texte collectif), « *Babel ou le choix du caviste : la bibliothèque à l'heure du numérique.* », in **Colloque virtuel « Text-e »** de la BPI Georges Pompidou. Année 2001. Hélas plus en ligne ... <http://www.text-e.org/>

=====T=W=I=T=T=E=R=====

==> **Mars 2006.** Lancement d'un site de micro-blogging baptisé Twitter. Son principe : la diffusion de messages de 140 caractères. Il rencontrera le succès qu'on lui connaît aujourd'hui.

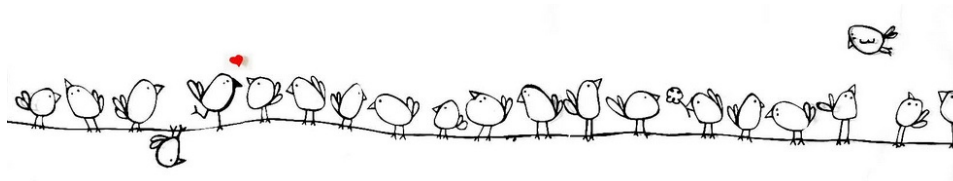
==> **Février 2010. Twitter "publie" chaque jour plus de 50 millions de "messages".**

==> **4 Mars 2010. Le cap des 10 milliards de tweets** est franchi.

=====T=W=I=T=T=E=R=====

**1er août 2006. Le – tardif – réveil.** La loi française étend à **Internet le principe du dépôt légal**. Toutes les bibliothèques s'y mettent avec plus ou moins de bonheur, plus ou moins de transparence.

**23 juin 2009 :** La BnF donne **quelques chiffres sur le dépôt légal Internet depuis 1996** : plus de 13 milliards de fichiers sont disponibles, organisés en collections mais **hélas uniquement consultables sur place et avec une accréditation** ... alors qu'au moins une partie du dépôt légal Internet de la bibliothèque du Congrès est **librement accessible et consultable en ligne**.



**"PATRIMOINE DU TEMPORAIRE".** La mission et le réflexe patrimonial des bibliothèques ou des archives a donc permis de relever le défi posé par le passage massif au numérique. Au-delà des seuls "documents numériques", la "sphère Internet" est également entrée dans un cycle de préservation à long terme. Il semble aujourd'hui évident pour tout le monde que la préservation d'une "archive de l'Internet" est au moins aussi importante que celle de manuscrits anciens ou des premières traces d'écriture. Il n'est pas inutile de rappeler qu'au

tout début des années 2000, en France, ce combat là paraissait encore... très loin d'être gagné ! Mais à l'échelle de l'histoire de la conservation des supports et des savoirs, le patrimoine du temporaire évoqué en 2001 n'aura mis que quelques années à trouver sa justification et son ancrage dans les pratiques de conservation de notre civilisation.

**14 avril 2010** : une **dépêche AFP** tombe, indiquant que la bibliothèque du Congrès (LoC) *"accueillerait les milliards de messages postés sur le site de microblogs Twitter depuis son lancement en mars 2006."* L'annonce est simultanément faite sur **le blog de la LoC**.

### **PATRIMOINE DU SUPERFLU. Pourquoi archiver la totalité de Twitter ???**

Les raisons de cet archivage peuvent être résumées en une phrase, extraite de la **Foire Aux Questions** mise en ligne très peu de temps après cette annonce par la LoC :

*"The Library of Congress collections include items such as the very first telegram ever sent, by telegraph inventor Samuel F.B. Morse, oral histories from veterans and ordinary citizens, and many other firsthand accounts of history."*

Du côté du premier public de cette nouvelle archive numérique, on trouve naturellement des historiens. **Historiens pour lesquels** :

*"Most of our sources are written after the fact, mediated by memory — sometimes false memory," Ms. Taylor said. "And newspapers are mediated by editors. Tweets take you right into the moment in a way that no other sources do. That's what is so exciting."*

**L'article du NYTimes** d'où est extraite la citation précédente fait également remonter des avis plus nuancés, notamment au regard des questions de "vie privée". Pour évacuer d'un seul trait l'ensemble des aspects polémiques autour de cette annonce, précisons que :

la LoC n'a rien payé (il s'agit d'un don de la société Twitter)  
que ne seront légués que les tweets des comptes publics  
qu'il y aura une barrière de six mois (minimum) entre la date de publication du tweet et son "versement" dans les archives de la LoC pour une exploitation ultérieure.  
Sur l'ensemble des questions précédentes, il vous faut à tout prix lire le billet de **Lionel Maurel sur le "patrimoine impossible" que représente cet archivage de Twitter**.

**Mémoires documentaires externalisées.** Depuis qu'elles existent, l'ensemble des techniques de "la documentation" ont eu pour objet de rendre d'abord plus efficace/efficente puis plus massive, systématique et transparente l'externalisation de nos mémoires documentaires. En décidant d'archiver l'ensemble de ce flot conversationnel qu'il faut bien qualifier de mémoire immédiate, la bibliothèque du Congrès fait faire un pas de plus à **l'archivistique** dans son ensemble. À tout le moins elle en étend le périmètre. A moins qu'elle ne le redéfinisse entièrement.

Pour bien comprendre les "vraies" raisons de cette conservation patrimoniale d'un nouvel ordre, d'une nouvelle nature, il faut lire le rapport final du *"Blue Ribbon Task Force on Sustainable Digital Preservation and Access"* (= groupe de travail sur les politiques publiques de conservation et d'accès numérique), groupe de travail à l'initiative duquel on trouve la NSF et la fondation Mellon en collaboration avec la Bibliothèque du Congrès, JISC, le CLIR et les Archives nationales US. Je m'en tiendrai aux passages synthétisés et traduits par Jean-Michel Salaün dans son billet **"Économie de la conservation numérique"** :

**1. La demande pour une conservation numérique est une demande dérivée.** Autrement dit, la demande n'est pas directe, on ne conserve pas pour conserver, mais pour donner accès à l'avenir à des informations numériques.

C'est bien ce qui s'est produit avec Twitter. La demande n'émane pas des créateurs du service, pas davantage que des usagers de la bibliothèque du Congrès.

**2. Les matériaux numériques sont des biens durables dépréciables.** *Un bien durable dépréciable est quelque chose qui dure longtemps en produisant de la valeur continuellement, mais la qualité et la quantité de cette production peut décliner si des actions ne sont pas engagées pour maintenir la viabilité ou la productivité du bien. (..)*

Twitter est certes un bien durable dépréciable... les deux derniers qualificatifs étant directement indexés sur sa capacité à trouver un modèle économique dans les prochains mois.

**3. Les biens numériques sont des biens non-rivaux et autorisent les passagers clandestins.** *Les biens numériques sont des biens non-rivaux, car il suffit qu'un acteur conserve un bien, il l'est pour toute intention ou objectifs conservé pour tous. Dans ces circonstances, l'incitation pour un seul acteur à assumer les coûts de la conservation est affaiblie, puisque les autres pourront profiter gratuitement des bénéfices.*

Twitter (la société) peut dès lors légitimement jouer les "passagers clandestins" puisqu'il n'aura pas à assumer le coût de la conservation ...

#### **4. La conservation numérique est un processus dynamique qui dépend du chemin suivi.**

Cette caractéristique est la plus originale et la plus spécifique au numérique et donc la plus intéressante. Dans l'analogique, le processus de conservation venait en fin du cycle de vie du bien, c'est-à-dire à la dernière étape du circuit classique de création-production-diffusion. Dans le numérique, chaque étape peut influencer sur le processus de conservation et celui-ci implique des décisions à chaque stade.

**Revenons-en maintenant à la question initialement posée : pourquoi archiver la totalité de Twitter ?** Aucune bibliothèque n'archive la totalité des sites web produits. Elle effectue un nécessaire travail de sélection : sites des événements sociaux ou politiques marquants (élections par exemple), sites de personnalités publiques importantes, etc., avec une difficulté supplémentaire dès que l'on sort de l'information "institutionnelle" : quels blogs archiver ? Selon quels critères ? En fonction de quels paramètres de choix ?

Il est deux raisons d'être à la science de l'archivistique : l'obligation et le choix. L'obligation d'archiver des supports arrivés à la fin de leur cycle de vie et le choix des "items documentaires" qui, parmi l'ensemble de ces supports en fin de cycle, sont dignes d'intérêt ou recouvrent une certaine "valeur".

Alors pourquoi archiver la totalité de Twitter quand on peut, dans les blogs, les forums, les sites web individuels, les "murs" Facebook, etc., quand on peut, disais-je, aller puiser partout ailleurs des éléments qui, pour reprendre l'argumentaire de la tribune des **historiens dans le NYTimes**, ne sont pas encore "médiés" par une activité mémorielle ?

**Les trois (vraies ?) raisons.** Il y a – à mon avis – trois raisons profondes à l'archivage de la totalité de Twitter ; trois raisons qui ne sont pas inscrites dans la foire aux questions dédiée. Trois raisons auxquelles j'en ajoute subrepticement une quatrième : le buzz :-). La LoC est passée maître dans les stratégies de communication en tout genre et il est clair que cet archivage du service médiatiquement le plus en vue actuellement aura, en termes d'image, de substantielles retombées.

**Primo : la simplicité.** Il est beaucoup plus simple de prendre, en vrac, la totalité de l'archive twitterienne que de l'investiguer pour y effectuer un travail de sélection en amont. On prend tout et on effectuera plus tard l'indispensable travail de tri pour isoler – par exemple – les tweets des personnalités politiques, tel **ce Tweet d'Obama au soir de son élection historique**.



**Deuxio : le graphe social.** L'archive de Twitter ainsi constituée permettra – à des chercheurs, sociologues, historiens – de reconstituer le graphe relationnel de chaque individu choisi. Ainsi on peut lire dans **l'article du NYTimes** que :

*"Each message is accompanied by some tidbits of supplemental information, like the number of followers that the author had at the time and how many users the author was following. While Mr. Cohen said it would be useful for a historian to know who the followers and the followed are, this information is not included in the Tweet itself."*

Il ne serait donc possible que de compter le nombre de comptes suivis ("following") ou de comptes suiveurs ("followers") pour un utilisateur donné. Sauf que. Sauf que c'est oublier un peu vite que le "vrai" graphe relationnel d'un utilisateur est également matérialisé à l'intérieur même de ses tweets, notamment grâce au symbole "@" qui, lorsqu'il est suivi d'un nom

d'utilisateur de Twitter, permet de s'adresser à lui directement. Ainsi, en épluchant à l'aide d'un algorithme les "@" de n'importe quel compte Twitter archivé, il est très facile de reconstituer son réseau relationnel, au moins dans son premier cercle de proximité (les gens auxquels on s'adresse le plus et/ou qui s'adressent le plus à nous). Donc c'est bien l'archivage "d'un" graphe social d'une petite partie de l'humanité connectée qui sera ainsi "en mémoire" à la bibliothèque du Congrès. Si ce n'est déjà fait, **j'insiste une nouvelle fois pour que vous alliez lire le billet de Lionel Maurel**, eu égard aux considérables questions de vie privée ainsi posés ... Par ailleurs, comme le démontre ReadWriteWeb, **le nombre d'informations "embarquées" dans un tweet n'est pas, loin s'en faut, réductible à une chaîne de 140 caractères ...**

**Tertio : l'archive elle-même.** Si la LoC s'intéresse à Twitter, un service qui, à l'échelle du web et indépendamment de son incontestable succès reste encore très jeune (moins de quatre ans d'existence) et dont rien ne permet de dire qu'il sera ou non pérenne sous sa forme actuelle (il peut être racheté, absorbé, décliné de manière radicalement différente à ce qu'il est actuellement, ou bien encore disparaître), si la LoC s'intéresse à Twitter, à *la totalité* de Twitter, c'est parce dans cette totalité transparaît l'essence même de l'objet documentaire idéal. Twitter comme un parangon de la documentation et de l'archivistique numérique. Voici pourquoi. Twitter a valeur de contexte et d'illustration. Contexte et illustration du monde (réel) dans lequel il s'inscrit et qu'il vient précisément documenter, dont il atteste de chacun des mouvements importants ou accessoires, un monde dont il est le témoignage, la valeur de preuve, la trace mémorielle immédiate. Il est en même temps un objet documentaire suffisant (il se suffit à lui-même, il peut-être analysé pour lui-même) et épuisant (le nombre de possibilités d'exploitations qu'il autorise ne peut être rapidement "épuisé"). Enfin, il est porteur de ramifications identifiables et externalisées : les liens contenus dans les tweets, qui, à leur tour, permettent de documenter à nouveau le cadre énonciatif dans lequel ils s'inscrivent.

**Et donc ?** Et donc il y a quelque chose que je n'arrive pas exactement à cerner – sinon j'aurai fait un billet beaucoup plus court ;-) - mais qui me gêne profondément dans cet archivage de Twitter, au-delà même du débat – déjà assez gênant – **sur la notion de confidentialité, de vie privée et de droit à l'oubli.** *"Les collections de la bibliothèque du Congrès regroupent des documents comme le tout premier télégramme envoyé, par l'inventeur du télégraphe Samuel F.B. Morse (...) et plein d'autres documents de première main"*. Oui mais le premier télégramme de Samuel Morse n'a rien à voir avec le premier tweet d'Obama. D'abord parce qu'à la différence de l'invention du télégraphe, Twitter n'est pas une rupture technologique. Ensuite parce que la valeur historique du premier télégramme de Samuel Morse lui a été conférée... par l'histoire ! (même s'il est effectivement permis de supposer que Samuel Morse, en l'envoyant, avait conscience du moment historique, mais bon on va pas non plus chipoter, sinon j'arriverai pas à finir ce billet)

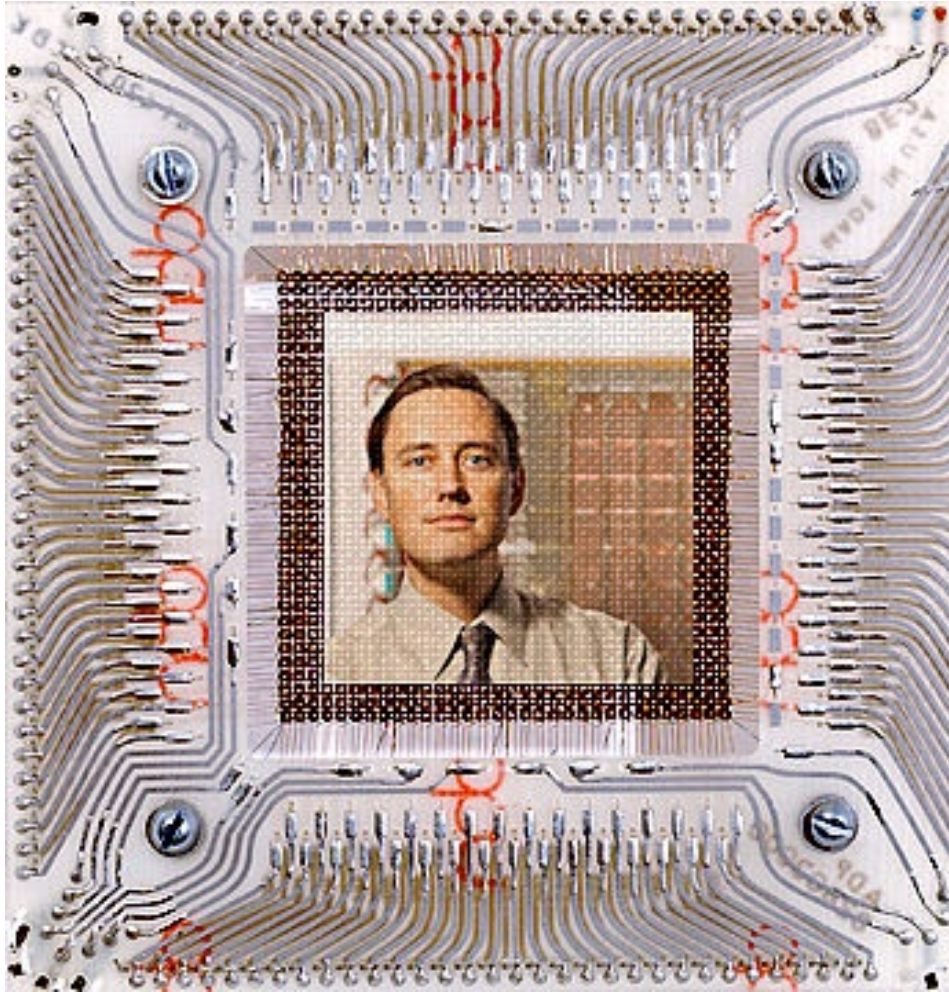
**CGU du 3ème type.** Ne faudrait-il pas désormais modifier en conséquence les CGU (conditions générales d'utilisation) de Twitter ? Mis à part des comptes "privés" (qui ne seront donc pas archivés), la nature d'un compte "public" me semble radicalement changer selon qu'il a vocation à être diffusé "publiquement sur le web" ou bien *"conservé six mois après sa publication et accédé indéfiniment dans le cadre des archives de la bibliothèque nationale des États-Unis à des fins d'étude et de recherche"*. Si c'est pour être disséqués dans les prochains siècles par des sociologues et historiens en tout genre, je vais peut-être hésiter à deux fois avant de poster des tweets sarcastiques ou graveleux sur l'actualité du Top 14 ou de la coupe d'Europe de rugby ...

**Par-delà.** Imaginez un instant les annonces suivantes. "La bibliothèque du Congrès annonce qu'elle archivera l'intégralité des comptes Facebook". Ou alors celle là : "La Bibliothèque nationale de France annonce qu'elle archivera l'intégralité des comptes publics Flickr et l'ensemble des photos qui y sont jointes." Et expliquez-moi en quoi des photos en disent plus sur nous que nos conversations sur Twitter ? Pourtant on peut aisément deviner le tollé que susciteraient de telles annonces.

**Avons-nous encore besoin d'archives publiques ?** Dans un monde où le cloud computing grignote le périmètre de l'archivistique, la réponse est naturellement oui. Oui parce que les archives publiques ont une vocation patrimoniale et pérenne qu'aucune société commerciale présente dans les nuages n'acceptera jamais d'assumer et/ou de porter. Oui mais non. Non si l'on se place un instant du côté de l'utilisateur lambda lequel est en permanence connecté à des archives, en permanence connecté à l'archive perpétuelle de sa propre documentation, de ses propres documents, et de ceux des autres, de tous les autres. La question doit peut-être être posée autrement.

**Avons-nous encore besoin de construire des archives** alors même que nous sommes en permanence immergés *dans* une archive immédiate, co-construite et suffisamment pérenne à notre échelle individuelle ? :

*“Par archive, j’entends d’abord la masse des choses dites dans une culture, conservées, valorisées, réutilisées, répétées et transformées. Bref toute cette masse verbale qui a été fabriquée par les hommes, investie dans leurs techniques et leurs institutions, et qui est tissée avec leur existence et leur histoire.” Michel Foucault, Sur l’archéologie des sciences (1968)*



**Mémoires. Cycles courts contre temps longs.** Avant, on “déclenchait” un processus d’archivistique documentaire soit au moment où un document parvenait en fin de cycle de vie, soit au moment où sa valeur historique était attestée ou signalée comme valant la peine d’être conservée. Avec l’archivage de Twitter, aucune de ces deux conditions n’est remplie. Or si l’on entreprend d’archiver de manière systématique, globale, des silos documentaires avant même qu’ils ne remplissent l’une ou l’autre de ces deux conditions, **on risque d’entrer dans un cycle de (re)production mémorielle inédit.**

On va produire de la mémoire sur de l’excès au lieu d’en produire pour répondre à un manque. On va créer de la mémoire sur de la mémoire. Alors que depuis des siècles, on avait appris à créer de la mémoire sur de l’oubli. Plus précisément, on avait appris à créer de la mémoire pour pallier l’oubli.

**Destin funeste de Funès.** Cette mémoire qui n’alimente qu’elle-même, cette hypermnésie dont souffrait le Funès de Borges, à l’échelle des institutions par nature depositaire de notre mémoire commune, se double d’une autre course folle : celle des individus eux mêmes englués dans des cycles mémoriels à la fois de plus en plus courts, de plus en plus instantanés, mais aussi de plus en plus denses, de plus en plus externalisés, de plus en plus rémanents. Après avoir inventé le droit à l’oubli numérique, le devoir d’inventaire de notre génération d’hypermnésiques compulsifs devra-t-il instaurer un simple droit au vide, un simple droit au non-dit, au non-inscrit, au non-rémanent ? Une législation de l’éphémère pour s’éviter les affres d’un patrimoine du temporaire ?

*“J’ai à moi seul plus de souvenirs que n’en peuvent avoir eu tous les hommes depuis que le monde est monde. Mes rêves sont comme votre veille. Ma mémoire, monsieur, est comme un tas d’ordure. (...) Il avait appris sans effort l’anglais, le français, le portugais, le latin. Je soupçonne cependant qu’il n’était pas très capable de penser. Penser c’est oublier des différences, c’est généraliser, abstraire. Dans le monde surchargé de Funès il n’y avait que des détails, presque immédiats.” Jorge-Luis Borges, Fictions – Funès ou la mémoire – Traduction P. Verdevoye ; Folio.*

**Le palimpseste interdit.** Après les technologies du faire, après celles de l’intelligence,

après celles de l'accès, **la période actuelle est celle des industries de la capillarité et de la percolation**. Industries dont l'archive est le nouveau cheval vapeur. Au-delà même de l'archivage de Twitter par la LoC, l'engrammation permanente risque d'empêcher à terme, le travail mémoriel authentique, lequel s'inscrit nécessairement dans la distance, dans le manque, et, pour partie, dans l'oubli préalable.

Le web se constitue est d'ores et déjà constitué comme un immense paratexte mémoriel que rien n'enfreint et que nul n'autorise ; un immense paratexte mémoriel que rien n'autorise et que nul n'enfreint ; et qui se déroule simplement. Un palimpseste sur lequel tout recouvrement est illusoire tant chaque strate reste indéfiniment accessible à qui le souhaite. La négation même du palimpseste comme recommencement possible.

A cela il me semble qu'il faut être vigilants. Non pour l'interdire, le dénoncer ou alimenter le courant d'un nouveau luddisme mémoriel. Mais pour... s'en souvenir le temps venu. **Nos vies, notre mémoire**.

—

Billet initialement publié sur **Affordance.info**

Illustrations CC Flickr **edu\_fon** ; **The Wren Design** ; **Desmond Kavanagh** ; **jurvetson** ; grande une **Trois Têtes (TT)**

## 1 ping

Les tweets qui mentionnent Twitter: un patrimoine superflu(x) ? » Article » owni.fr, digital journalism -- Topsy.com le 12 mai 2010 - 16:23

*[...] Ce billet était mentionné sur Twitter par Aurélien Fache, Georges-Ed. Legré et Romain LABOUREE, Flash Presse. Flash Presse a dit: Twitter: un patrimoine superflu(x) ? <http://goo.gl/fb/gqNW3> [...]*