

# TWITTER: L'IMPOSSIBLE ARCHIVAGE ?

LE 12 MAI 2010 LIONEL MAUREL (CALIMAQ)

L'archivage du service par la Bibliothèque du Congrès soulève des questions sur le respect de la vie privée. Les conditions d'accès imposées posent aussi problème. L'enjeu derrière, c'est l'accès à l'information et la préservation du domaine public.

La nouvelle avait fait sensation, il y a quinze jours : **Twitter annonçait** faire don de l'intégralité de ses archives à la Bibliothèque du Congrès à des fins de conservation et de recherche et c'est, comme il se doit, **par un tweet** que la LoC (Library of Congress) célébrait l'entrée du rétrospectif de Twitter dans ses collections.

Soit quand même **plus de dix milliards de gazouillis !**

Le même jour, **Twitter et Google annonçaient** le lancement d'un nouveau service du moteur de recherche – Google Replay – qui permettra d'effectuer des recherches chronologiques au sein du contenu de Twitter, avec à terme l'intention de balayer l'intégralité des archives, depuis son lancement en 2006.

Cette "patrimonialisation éclair" de Twitter, consacrée par son entrée à la Bibliothèque, a quelque chose qui peut surprendre, même si elle s'accorde bien avec le tempo ultra-rapide du microblogging. On peut aussi penser que Twitter contient déjà de vrais petits morceaux d'histoire, comme les tweets liés au hashtag **#iranelections, le pépiement de la victoire électorale de Barack Obama** ou encore **le premier tweet envoyé depuis l'espace !** La globalité des tweets forme aussi une formidable archive du quotidien, ouvrant **de nouveaux champs à la recherche historique**, par le biais du *data mining*.

Pour autant, les commentaires français n'ont pas manqué de relever que ce transfert des archives de Twitter à la Bibliothèque du Congrès soulevait également des questions, à propos **du respect de la vie privée** ou **du droit d'auteur**. Et de l'autre côté de l'Atlantique, où l'on ne plaisante pas avec la protection de la *privacy*, c'est une véritable polémique qui s'est déclenchée, plaçant la LoC dans une position assez délicate (vous en trouverez **quelques échos ici**, rassemblés par le blog ArchivesNext).

À tel point que la responsable de l'archivage numérique de la LoC **a donné une interview** pour apporter des précisions sur le partenariat avec Twitter. La Bibliothèque a également **publié cette semaine une FAQ**, accompagnée (et c'est là que ça devient intéressant !) de l'accord par lequel Twitter a fait don de ses archives (**ici**).

L'analyse de ce document soulève des questions assez troublantes, qui me laissent penser que la constitution de la mémoire numérique, dans le cadre de partenariats public-privé, est loin encore d'avoir trouvé son assiette juridique.



Il faut en effet garder à l'esprit que Twitter est l'exemple type d'un service web 2.0, dont le contenu est produit par ses utilisateurs (**User Generated Content**). La question de la propriété de ces contenus est **loin d'être aisée à appréhender**, surtout que les micromessages de Twitter possèdent une nature assez particulière.

## Copyright or not ?

L'été dernier, j'avais écrit une série de billets dans lesquels j'avais essayé de montrer que les tweets échappent dans leur immense majorité à la propriété intellectuelle, par manque d'originalité ou de mise en forme (**ici** ou **là**). Dès lors, ils relèvent plutôt du statut de l'information brute ou des données, et ne devraient pas pouvoir faire l'objet d'une appropriation (y compris d'ailleurs par leurs propres "auteurs"!).

Pourtant, quand on lit le Gift Agreement publié par la Bibliothèque du Congrès, on se rend compte que c'est sur le copyright que Twitter s'appuie pour délivrer à la LoC une licence d'utilisation de ses contenus :



***2) Copyright : Donor grant an irrevocable nonexclusive licence to the library for such rights as the Donor has the right to transfer or licence under the Twitter Terms of Service in place at the time of the gift or before. The current, as of the effective date, and previous Terms of Service are appended.***



Twitter utilise en effet ses **CGU (Conditions Générales d'Utilisation)** pour revendiquer un droit d'auteur (copyright) sur les contenus produits par ses usagers :



***You retain your rights to any Content you submit, post or display on or through the Services. By submitting, posting or displaying Content on or through the Services, you grant us a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such Content in any and all media or distribution methods***

*(now known or later developed).*



C'est cette licence – *royalty-free license (with the right to sublicense)* – que Twitter met en œuvre dans sa convention de don à la LoC et qui lui confère juridiquement sa qualité de *Donor*.

Le problème, c'est qu'on peut sérieusement se demander si cette chaîne de concessions est valide et comment Twitter peut revendiquer un droit sur ses contenus, si les utilisateurs eux-mêmes ne possèdent pas réellement de droit d'auteur sur leurs propres tweets...

C'est une aporie que j'avais **déjà relevée l'année dernière**, lorsque les CGU de Twitter ont changé et que le petit oiseau bleu est subitement devenu plus agressif vis-à-vis de ses utilisateurs.

En effet, à l'origine (souvenez-vous !), Twitter ne revendiquait aucun droit de propriété intellectuelle sur ses contenus et incitait même ses utilisateurs à les verser dans le domaine public :



#### ***Copyright (What's Yours is Yours)***

***1. We claim no intellectual property rights over the material you provide to the Twitter service. Your profile and materials uploaded remain yours. You can remove your profile at any time by deleting your account. This will also remove any text and images you have stored in the system.***

***2. We encourage users to contribute their creations to the public domain or consider progressive licensing terms.***



J'estime pour ma part que les nouvelles CGU de Twitter n'ont pu avoir pour effet de changer la nature de son contenu, et que celui appartenait et appartient toujours, dans sa plus grande partie, au domaine public.

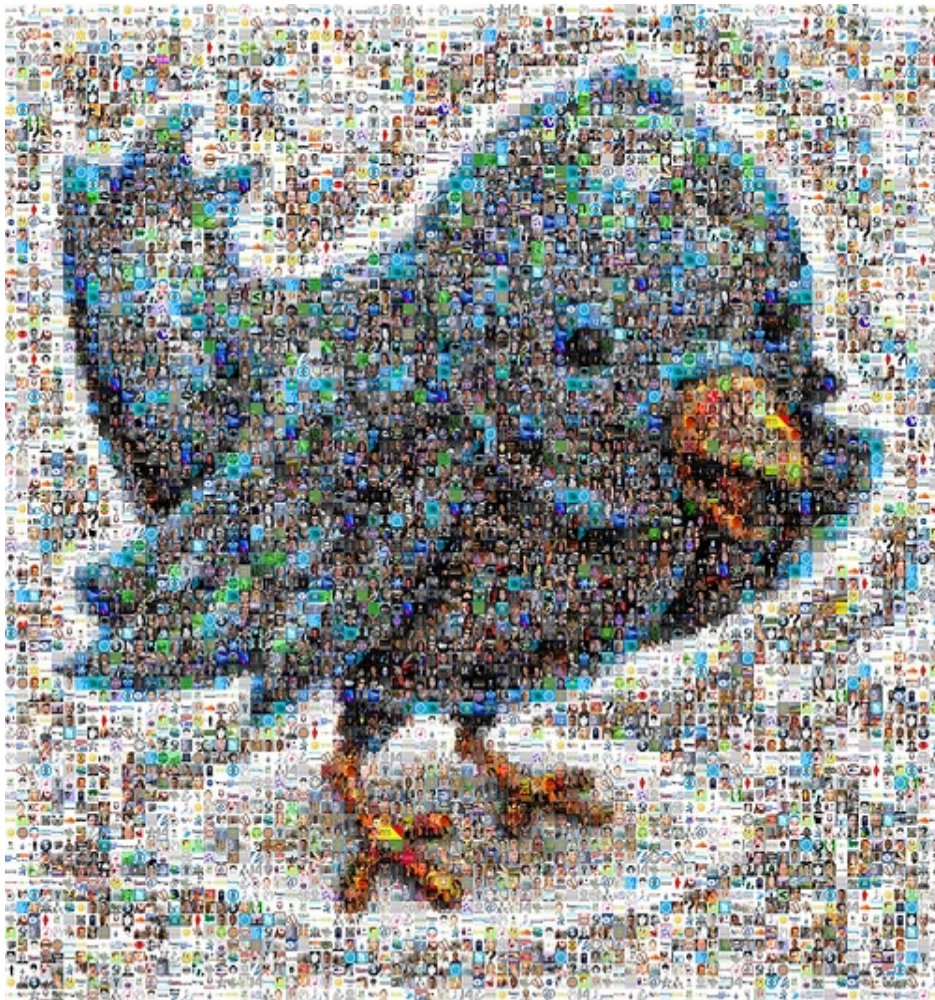
Dès lors, le fondement même de la convention de don à la LoC me paraît douteux, et même préjudiciable, puisqu'il entérine le geste d'appropriation que Twitter a opéré vis-à-vis des contenus produits par ses utilisateurs, avec l'aval de la Bibliothèque.

**Certaines analyses** estiment que Twitter a manqué aux obligations qui le liaient avec ses utilisateurs, mais que la Bibliothèque, de son côté, n'avait nullement besoin de cette convention de don pour archiver le contenu de Twitter. La loi américaine sur le dépôt légal permet en effet à la LoC de collecter toutes les formes de publication. Elle dispose déjà d'ailleurs **d'une archive du web**, constituée par des captures de sites.

## **Patrimoine public contre vie privée ?**

D'autres voix se sont élevées aux États-Unis pour contester à Twitter la faculté de donner ainsi ses archives à la Bibliothèque, mais cette fois au nom du respect de la vie privée (**ici, là ou là**).





Fred Stutzman, sur son blog, se livre ainsi à une intéressante interprétation restrictive des CGU de Twitter :



*[...] as long as your content is on Twitter, Twitter can do what they want with it. Fine. But what if you remove your content from Twitter? Wouldn't Twitter's licensing of your content to the LoC also expire? Twitter needs to address exactly how we can pull our content out of the archive when we want.*

*A broader question is why Twitter didn't just build this as an opt-in service. Or even, less preferably, an opt-out service. Is the collection so important that it is worth compromising user privacy ? I've got a feeling that there are certain assumptions around "public" content and the feel-good vibe of the Library of Congress that led to a lack of critical thinking about the implications of this move.*



On touche ici à des questions qui mêlent à la fois la propriété sur les contenus et le respect de la vie privée, d'une manière qui fait penser au **débat sur le droit à l'oubli** ayant lieu actuellement en France.

Beaucoup d'utilisateurs américains se sont inquiétés par exemple que les tweets de comptes privés puissent être consultables à la LoC, de la même façon que les **Direct Messages**, qui ne sont pas réellement "publiés sur Twitter", au sens de "rendus publics". La question reste entière également pour les données de géolocalisation. Il semblerait par contre que les informations des profils personnels, ainsi que les tweets annulés ne feront pas partie de l'archive.

La question de l'opt-out est aussi déterminante : les utilisateurs qui annulent leurs comptes

sur Twitter pourront-ils aussi exiger le retrait de leur message de l'archive consultable à la LoC ? Pourront-ils effectuer des retraits ciblés de certains de leurs messages seulement et comment ?

**Dans son interview**, la responsable de l'archivage numérique de la Bibliothèque laisse entendre que cet opt-out est une question qui concerne Twitter et ses usagers et pour laquelle la Bibliothèque ne souhaite pas jouer le rôle d'un intermédiaire. Elle indique également que la Bibliothèque envisage d'anonymiser les contenus pour éviter toute atteinte au respect de la vie privée, travail que l'on imagine titanesque sur 10 milliards de tweets !

## Un don asymétrique

Un autre aspect de la convention de don qui me paraît encore plus contestable est celui des conditions d'accès qui ont été imposées par Twitter à la LoC.

Une période d'embargo de six mois a été instaurée avant que la Bibliothèque puisse donner accès aux contenus déposés par Twitter. La justification de ce délai n'est pas des plus claires. On imagine peut-être que Twitter ne souhaite pas que l'archive de la LoC devienne un Twitter bis. Cela dit, il y a **tellement de services tiers** qui reprennent ses contenus par le biais de son API que cette précaution paraît un peu dérisoire. On peut lire ailleurs que ces six mois permettraient aux utilisateurs de procéder au retrait de certains de leurs tweets sur leurs comptes, pour mettre en œuvre une forme d'opt-out.

Des restrictions d'usages assez fortes figurent également dans la convention : l'archive pourra être utilisée en interne par la Bibliothèque pour ses propres besoins et un accès pourra être donné dans ses emprises aux chercheurs habilités, après avoir signé un engagement de ne pas faire un usage commercial des contenus. La LoC ne pourra pas redistribuer l'archive à des tiers dans sa totalité ou une partie substantielle de celle-ci (**même pas à Hathi Trust ?**).

L'accord permet cependant à la LoC de diffuser sur son site, au terme des six mois d'embargo, des éléments de l'archive, à condition de bloquer l'indexation par les moteurs de recherche (**robots.text file**) et d'empêcher le téléchargement substantiel de contenus.

À lecture de ces conditions, j'ai éprouvé une sensation très désagréable, car elle me rappelle beaucoup celles que **Google impose à ses bibliothèques partenaires** dans le cadre du programme Google Books, notamment l'exclusivité d'indexation. Cela signifie que, même accessible en ligne, cette archive de la Bibliothèque restera quasiment invisible sur le web.

Et le parallèle est d'autant plus saisissant que dans le même temps, Google a effectivement mis en place, certainement au terme d'un accord avec Twitter, **son propre service de recherche dans les archives des tweets, Google Replay**. Celui-ci sera entièrement accessible en ligne, sans la restriction des six mois d'embargo, et bien entendu, il figurera dans les résultats du moteur (et dans ceux de ses concurrents ?). Sans compter que Google n'a pas l'air de beaucoup s'embarrasser de son côté des questions de protection des données personnelles...

Encore une fois, on voit naître une situation asymétrique entre l'usage que les institutions publiques pourront faire des contenus, alors même qu'elles en assureront la préservation à très long terme, et l'usage (devrais-je dire l'exploitation...) que pourront en faire les acteurs privés. Ici encore, j'ai envie de parler d'une forme d'**eugénisme documentaire** qui s'organise par le biais de telles restrictions contractuelles, et ce (et c'est le plus grave !) avec l'accord des bibliothèques.



Encore et toujours, le véritable enjeu, c'est l'accès à l'information et la préservation du domaine public. Comme je l'ai dit, les tweets ne sont pas dans leur immense majorité des contenus protégés par le droit d'auteur. Il s'agit d'informations et de données qui ne devraient pas faire l'objet de telles restrictions. Twitter d'ailleurs ne possède certainement pas de titre valide pour imposer ces limitations à la LoC, si ce n'est celui qu'il tire, de manière fort douteuse, de ses CGU.

Et le pire, c'est que les restrictions imposées par la convention de don sont valables perpétuellement, sans limitation de durée...

Rien n'est plus faux que de dire que cet accord entre Twitter et la LoC permet de constituer ses archives en un patrimoine qui pourra être préservé pour le futur.

C'est tout le contraire qui se produit : ce qui est préservé pour l'éternité, c'est l'acte d'appropriation que Twitter a accompli l'année dernière sur les contenus produits par ses usagers. Ces contenus seront peut-être conservés dans leur intégrité par la LoC, mais ils ne pourront jamais regagner le domaine public.

C'est exactement la même contradiction qui entache les accords signés entre les bibliothèques et Google à propos de la numérisation du patrimoine écrit.

Pour ma part, j'avais choisi de placer tous mes tweets sous **la licence Creative Commons Zéro (CC0)**, grâce à l'application **TweetCC**. Cette licence me permet de certifier que je renonce à tous mes droits sur mes tweets et que je les verse au domaine public, sans aucune restriction.

**C'est par un tweet** que j'ai manifestée cette volonté, à laquelle j'accorde beaucoup d'importance.

Ce tweet est désormais quelque part enfoui dans l'archive de la LoC... éternelle lettre morte ...

—

Billet initialement publié sur **:: S.I.Lex ::** sous le titre "Twitter archivé à la Bibliothèque du Congrès : un patrimoine impossible ?"

Photo CC Flickr **bbluesman**

**NEO**

le 10 février 2011 - 18:47 &bullet; SIGNALER UN ABUS - PERMALINK



*amanda hurley curves magazine  
wiring a subpanel box  
common orcish translate  
three actions of prostaglandins  
dead porsche girl*



videos morras de secu perreando  
jawatan kosong swasta indeks  
disney com thehub  
best hair salons in pasadena ca  
kwentong bikol  
delicious magazine jerk chicken recipe  
napa 1056 oil filter  
300 rum maximum effective range  
relatos de incesto  
discount universal soul circus tickets  
www hinh tu the sex  
circumference diameter powerpoint  
teacup chihuahua tx hillcountry  
celebrita sotto la gonna  
culiacan plebitas

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

## 2 pings

Les tweets qui mentionnent Twitter: l'impossible archivage ? » Article » owni.fr, digital journalism -- Topsy.com le 12 mai 2010 - 15:43

[...] Ce billet était mentionné sur Twitter par Aurélien Fache, Owni, Alexis Jaillet, Mickael Mingot, Badibule et des autres. Badibule a dit: Twitter: l'impossible archivage ? <http://bit.ly/cr7CcY> [...]

20 raisons qui font que vous n'engendrez pas assez de followers sur Twitter | le 16 août 2010 - 19:30

[...] Source de l'image : Owni.fr [...]