

# FACEBOOK, LA VIE TÉLÉCHARGÉE

LE 18 NOVEMBRE 2010 OLIVIER ERTZSCHEID

Saisissant l'opportunité offerte par Facebook, Olivier Erztscheid a téléchargé l'ensemble de ses données personnelles. Il y voit l'émergence d'un nouveau langage documentaire.

Ca y est c'est fait. Sur mon compte Facebook, dans l'onglet "Account" > "Account settings", j'ai vu apparaître le petit lien magique que j'attendais tant : "Download your information"

<b>Privacy</b>	<a href="#">manage</a>
Control what information you share.	
<b>Account Security</b>	<a href="#">change</a>
<b>Download Your Information</b>	<a href="#">learn more</a>
<b>Deactivate Account</b>	<a href="#">deactivate</a>

J'allais enfin pouvoir télécharger toutes mes données personnelles. Après avoir indiqué que "oui, oui, je suis sûr de bien vouloir les télécharger", et une heure plus tard, j'ai reçu dans ma boîte mail un message de Facebook avec un lien d'activation pour accéder au précieux fichier.

Mais il m'a d'abord fallu doublement montrer patte blanche : une première fois en redonnant mon mot de passe (ok, précaution élémentaire), et une deuxième fois en jouant au jeu des photos de mes amis.

## Download Your Information

Get a copy of the data you've put on Facebook.

This is a copy of all of the personal information you've shared on Facebook. In order to protect your information, we will ask for authentication to verify your identity. [Learn More](#) about downloading a copy of your information.

**WARNING:** This file contains sensitive information. Because this download contains your profile information, you should keep it secure and take precautions when storing, sending or uploading it to any other services.

### Security Check

To verify that you are the owner of this account, please identify the people tagged in the following photos. If you aren't sure about a question, please choose "I'm not sure."

[Start](#)

Je traduis :

“

***Pour vérifier que vous êtes bien le propriétaire de ce compte, merci de reconnaître les personnes tagguées sur ces photos suivantes.***

”

Il faut reconnaître 5 amis et on l'on n'a droit qu'à 3 erreurs ("I'm not sure"), soit 8 questions en

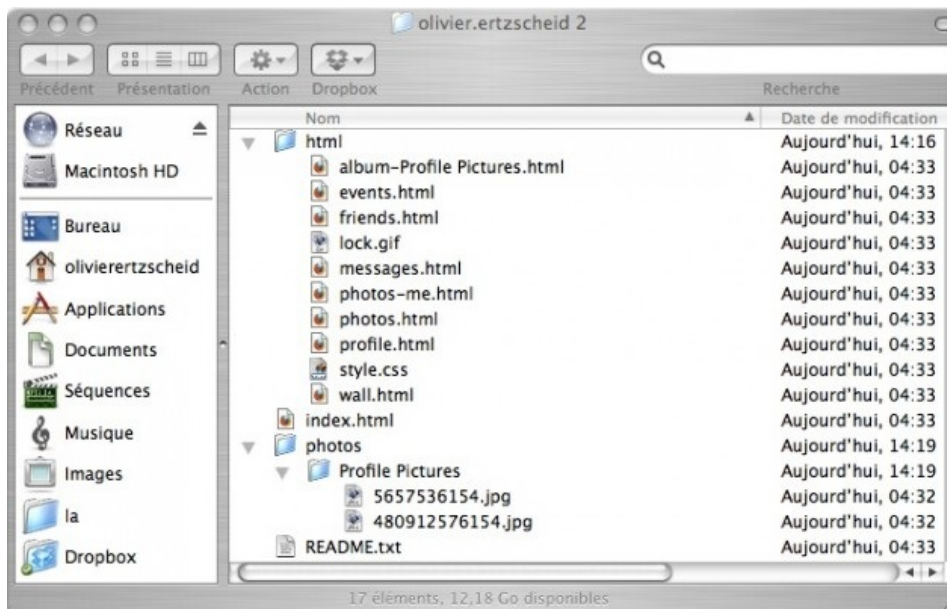
tout. Le problème c'est que les photos affichées ne sont pas des *portraits* de vos amis, mais des photos prises au hasard dans le photostream de tous vos contacts Facebook, c'est à dire toutes les photos déposées par tous vos amis.

Et là première angoisse : je suis ami avec plein de gens, mais surtout j'ai plein "d'amis d'amis" et encore davantage "d'amis étudiant(e)s". Et autant vous dire que je ne passe pas mon temps à regarder toutes les photos postées sur Facebook par tous mes "amis". Et donc ça ne loupe pas, on me demande de reconnaître quel est l'ami qui se cache derrière ces photos :



Voilà voilà voilà ... Donc ben du coup je n'ai plus droit qu'à 2 erreurs en espérant que les autres photos seront plus ... explicites. Quelques reconnaissances plus tard, c'est gagné :

Une fois le téléchargement effectué, je me retrouve avec ça :



Le dossier avec :

mes photos de profil (album-Profile Pictures.html)  
la liste alphabétique de tous mes amis (friends.html)  
les événements auxquels j'ai été convié (events.html)  
mes courriers reçus sur facebook (messages.html)  
tout ce que j'ai posté sur mon "mur" (profile.html), soit – en ce qui me concerne – l'équivalent de 150 pages Word, et ce qui m'a permis de retrouver la date précise de mon arrivée sur Facebook: c'était le 11 Juillet 2007 à 13h40 et mon premier statut disait: "At home". Largement de quoi **imprimer un aussi épais que futile "egobook"**.

Le dossier avec :

toutes mes photos (c'est à dire en ce qui me concerne, uniquement 2 pauvres photos de profil)  
Soit un dossier complet d'1,3 Méga-octets.

**Remarque :** ce dossier et les fichiers qu'il contient est un "à plat". Ainsi, dans le fichier "profile.html" on retrouve bien – et on peut activer – les liens hypertextes postés sur mon mur ou dans mes "statuts", mais on ne retrouve naturellement pas le graphe relationnel/navigationnel qui constitue le vrai trésor de guerre de Facebook, notamment – depuis l'activation de la fonctionnalité – les personnes qui ont "aimé" (= le fameux bouton "like") tel ou tel statut.

## Et maintenant un peu de maths

Sachant qu'un individu (moi en l'occurrence) peut être considéré comme un utilisateur type du réseau social Facebook, et sachant qu'il y a au moins 500 millions d'utilisateurs sur Facebook, quel est le poids des données personnelles détenues par Facebook ?

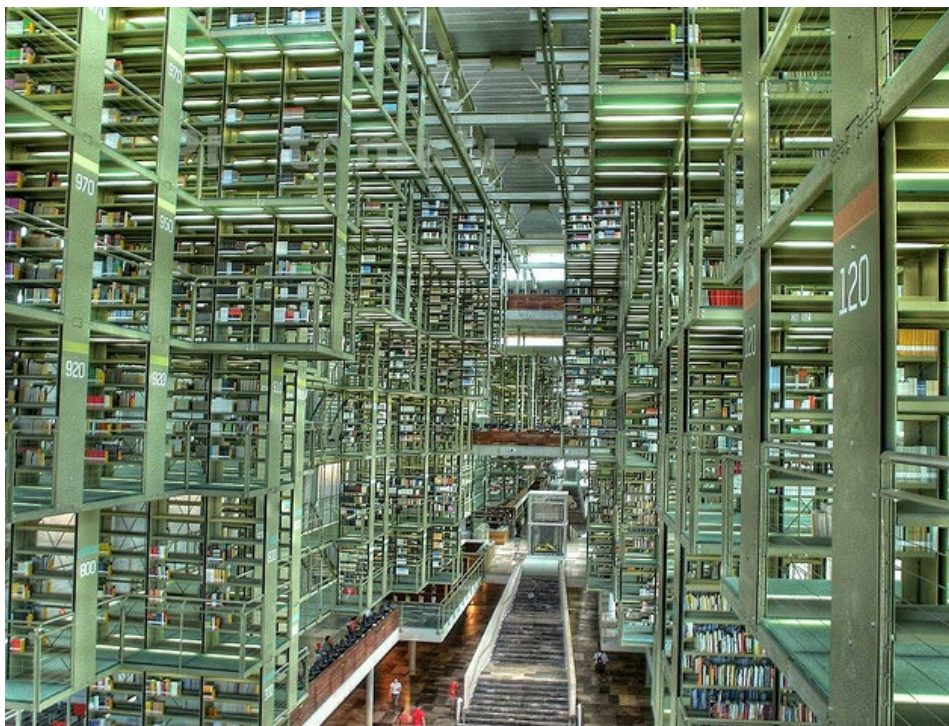
*Solution :* 500 millions multiplié par 1,3 méga-octets = 650 millions de mégaoctets.

**Ce qui nous donne :** 650 Téra-octets de données personnelles disponibles sur Facebook  
*[EDIT: selon le commentaire d'un employé de Facebook posté sur le blog d'Olivier Ertzscheid, ce chiffre serait encore plus élevé, notamment parce que l'entreprise de Mark Zuckerberg a collecté plus de 160 téraoctets de photos sur la seule année 2007. Par ailleurs, Twitter stockerait pas moins de 4 pétaoctets (4.096 téraoctets) par an.]*

## 65 fois la bibliothèque du Congrès

Sachant que "10 téraoctets pourraient contenir toute la collection des ouvrages imprimés de la bibliothèque du Congrès" (**source**), on peut donc supposer que Facebook détient au moins l'équivalent de 65 bibliothèques du Congrès uniquement composées de données personnelles. Celle-ci comptant **plus de 33 millions d'ouvrages**, cela fait donc l'équivalent de (33 x 65) : 2.145 millions de livres de données profilaires.

Et encore, je dis bien "au moins" car dans les profils Facebook ce sont les photos qui prennent le plus de place, que je n'en ai que deux (photos) dans le mien et qu'en moyenne mes amis sont plus proches d'une bonne cinquantaine (de photos toujours), et je ne vous parle même pas de mes "amis-étudiants" (il n'est pas rare de voir plus de 300 photos dans certains profils). Mais bon on va pas chipoter, l'ordre de grandeur me semble déjà suffisamment parlant ...



## Mon mur, ma bataille

L'équivalent de 150 pages Word pour moi en 3 ans, et pour 499.999.999 autres, 2.145 millions de livres de données profilaires dont probablement plusieurs centaines de millions également et uniquement remplis de ces traces profilaires, conversationnelles, le reste étant



occupé par la documentation iconographique rétrospective des 500 millions d'habitants de cette communauté. Vertigineux.

## Facebook en grand ordonnateur du nouvel ordre documentaire mondial?

Facebook est, par le nombre, la **première communauté humaine** de la planète numérique. Si **l'homme est un document comme les autres**, et si l'ordre documentaire du 21<sup>ème</sup> siècle sera celui d'un pan-catalogue des individualités humaines, Facebook est en bonne place pour remporter la mise ou pour en devenir à tout le moins le grand ordonnateur, **le grand sachem de ce qui est "su"**.

## De la thésaurisation des profils au trésor de guerre

Au-delà des chiffres et des questions de "vie privée", il faut relire **ce billet de Tim Berners Lee**, une nouvelle fois visionnaire, dans lequel il évoque le **Giant Global Graph**.

Le net est un graphe d'ordinateurs connectés.

Le web est un graphe de contenus connectés, dont Google est, pour l'instant, l'outil qui permet le mieux d'en sonder les profondeurs, d'en donner l'image la plus "complète" possible.

Les réseaux sociaux sont un graphe d'individus connectés, dont Facebook est, pour l'instant, l'outil qui permet le mieux d'en sonder les profondeurs, d'en donner l'image la plus "complète possible".

L'enjeu est désormais de savoir qui sera le premier à pouvoir réunir la puissance des 2 graphes. La guerre des graphes a désormais officiellement commencée. Et comme toutes les guerres, elle fut d'abord larvée, chacun essayant de circonscrire au mieux ses frontières naturelles (les contenus pour Google, les profils pour Facebook).

Elle fut ensuite une série de petites offensives permettant de jauger les forces et faiblesses de l'adversaire en envoyant une petite armée le combattre sur son terrain : ce que tenta de faire Google en lançant "son" réseau social (Orkut), ce que tenta de faire Facebook en nouant alliance avec Microsoft pour s'installer – mais en restant dans ses frontières – sur le marché du "search".



## "Projet Titan"

Et puis un jour, à l'occasion d'une escarmouche, on sort l'artillerie lourde et on engage "officiellement" le début des hostilités. C'est désormais chose faite entre Google et Facebook.

L'escarmouche ce fut ce "cheval de troie" qui permettait à un nouvel arrivant sur Facebook d'importer rapidement l'ensemble de ses contacts Gmail pour densifier rapidement son réseau d'amis. **Jusqu'à ce que Google exige une contrepartie**, c'est à dire que les utilisateurs de Gmail puissent récupérer et importer – par exemple – leurs contacts Facebook. Un blocage par ailleurs **rapidement contourné par Facebook**.

Et l'on apprend (**sur Techcrunch**) que Facebook lancerait (demain lundi ?) son webmail, nom de code "projet Titan" (sic) :

“

***Le réseau social proposerait ainsi à ses utilisateurs, qui sont plus de 500 millions dans le monde, une adresse email personnelle @facebook.com, permettant d'envoyer des mails à tous les internautes, qu'ils aient une messagerie Hotmail, Yahoo ou Gmail. Aujourd'hui, les membres de Facebook ne peuvent envoyer de messages qu'aux autres membres du site.***

—  
(Source)

”

Le courriel et les webmails sont, pour ces deux acteurs, un cheval de Troie idéal et hautement stratégique permettant d'assiéger la place-forte de nos pratiques connectées :

D'abord parce que les mails rassemblent, parce qu'ils "synthétisent" notre réseau relationnel (nos "contacts"),  
Ensuite parce qu'ils sont le point d'entrée le plus aisé vers le cloud computing, le stockage "dans les nuages" ou plus exactement sur les serveurs de ces grandes compagnies de pans entiers de nos vies sociales (pièces jointes, documents de travail, photos, etc ...)  
Parce qu'ils contiennent également ce que nous avons à raconter de plus "intime", de plus "personnel" et qu'ils permettent donc d'affiner encore l'affichage de publicités "contextuelles" en scannant le contenu desdits mails,

Enfin parce que les webmails peuvent, au sein d'un écosystème semi-fermé (comme Facebook) ou semi-ouvert (comme Google et sa galaxie de services), constituer un point pivot autour duquel hiérarchiser l'ensemble des autres données affiliées à notre profil, à notre "empreinte numérique".

Techcrunch souligne ainsi que "*Facebook aurait les moyens à la fois de hiérarchiser les courriels et de les intégrer à ses autres fonctions (partage de photos, calendrier etc.) d'une façon très convaincante.*" (source)



## **Thesaurus : From society of query to society of contact**

Google est depuis longtemps emblématique de ce que Geert Lovink appelle une "**Société de la requête**". Facebook représente lui, les promesses d'une société des contacts étendus, ou distendus. Or requêtes comme contacts représentent la conquête et l'apprentissage d'un langage commun à l'humanité tout entière ; plus précisément un **thesaurus**, c'est à dire :



**Un type de langage documentaire qui consiste en une liste de termes sur un domaine de connaissances, reliés entre eux par des relations synonymiques, hiérarchiques et associatives.**



En l'occurrence, de l'approche commune de Google et Facebook on pourrait dire qu'elle vise à faire émerger et à circonscrire un type de langage documentaire qui consiste, en une liste de requêtes (déposées sur le moteur ou sur chacun de nos "murs"), portant sur l'ensemble des domaines de connaissance existants (des plus fondamentaux aux plus futiles).

Ces requêtes sont reliées entre elles par des relations lexicales (ingénierie linguistique) et associatives "amicales" ou "contactuelles" (ingénierie de la recommandation reposant sur des graphes relationnels) permettant de hiérarchiser l'ensemble des requêtes et des profils en fonction du contexte de la requête et/ou du profil du requêtant.

**Moralité.** L'étymologie du mot thésaurus désigne, en latin, le "trésor". De ce trésor là nous n'avons pas encore fini de mesurer la valeur. De ce trésor là nous connaissons déjà ceux qui veulent en être les grands avaleurs.

**Cet article a initialement été publié sur Affordance**

Crédits photo: Flickr CC [dbking](#), [Eneas](#), [Mosman Library](#), [malias](#)

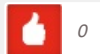
#### STEPHANE

le 18 novembre 2010 - 16:54 &bullet; SIGNALER UN ABUS - PERMALINK



*Une chose que j'ai remarque en essayant cette fonction c'est qu'il est possible de remonter en arriere de plusieurs annees... Le plus vieux post dans cette archive datait d'il y a 4 ans. Il me semblait pourtant que l'UE avait legifere sur la duree de conservation des donnees?*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

#### RENO

le 18 novembre 2010 - 17:16 &bullet; SIGNALER UN ABUS - PERMALINK



Hello,

*Disclaimer : je travaille chez FB mais ce que je dis n'engage que moi :)*

*Article intéressant.. mais vous êtes (très très) loin du compte avec les 650To. FB a des dizaines/centaines de To de nouvelles données chaque \*jour\*.*

*Voici quelques exemples d'infos publiques là dessus :*

*<http://blog.facebook.com/blog.php?post=2406207130>*

*(160To rien que de photos en 2007, avec juste quelques dizaines de millions d'utilisateurs)*

*<http://www.scribd.com/doc/22847003/Facebook%E2%80%99s-Petabyte-Scale-Data-Warehouse-Using-Hive-and-Hadoop>*

*Autre point de comparaison : Twitter, qui est beaucoup, beaucoup plus petit que FB, stocke 4Po (4096 To) de tweets par an, sans photos ni vidéos !*

*<http://www.neowin.net/news/storing-tweets-requires-four-petabytes-of-data-a-year>*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

PAUL



le 18 novembre 2010 - 19:05 &bullet; SIGNALER UN ABUS - PERMALINK



*J'ai pas encore l'option... C'est pas dispo chez tout le monde ?*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### GAGANAUSAURE

le 18 novembre 2010 - 23:00 &bullet; SIGNALER UN ABUS - PERMALINK



*On arrête pas le progrès !  
Vive la STASI, ludique, citoyenne et capitaliste...  
Les responsables de toutes les agences de renseignements du monde n'osaient pas rêver une telle ardeur et autant d'honorables correspondant !*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### XIAOXIN

le 19 novembre 2010 - 10:22 &bullet; SIGNALER UN ABUS - PERMALINK



*@Paul :  
il me semble que j'avais lu que l'option de téléchargement des données personnelles ne se trouvait que dans l'interface anglaise de facebook ... si on est en français il faut passer en anglais (peut-être US)*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### TUEURSNET

le 19 novembre 2010 - 17:20 &bullet; SIGNALER UN ABUS - PERMALINK



*Vous croyez y être, mais vous n'y êtes pas. Vous ne croyez pas y être, mais si, vous y êtes. Internet est le lieu de toutes les médisances, nuisances et vengeances en tous genres. A l'heure du numérique, les pendules juridiques n'indiquent toujours pas la même heure. Décalage qui peut vous couter très cher, si vous vous baladez en étant persuadé qu'il n'y a aucune raison pour que quelqu'un vous plante un couteau dans le dos... C'est pourtant ce qui arrive tous les jours sous nos yeux parce que les visages sont masqués et les messages non signés. Tout est virtuel, à part les larmes ou le sang, qu'ils font couler... Vous pouvez toujours faire appel à des nettoyeurs sur le web, qui peuvent riposter en menaçant, en noyant ou en étouffant vos plus proches ennemis. Mais le plus court chemin reste : la bombe ou la tombe ?  
<http://www.tueursnet.com/index.php?journal=Balle%20de%20Webmaster>*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### DFITZ

le 21 novembre 2010 - 10:49 &bullet; SIGNALER UN ABUS - PERMALINK



*@Paul, je te confirme que pour accéder à cette fonctionnalité, tu dois modifier la langue de l'interface et choisir english (us), ensuite tu reviens à la gestion du compte et tu verras apparaître le lien.*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

### R\_GRANDMORIN

le 22 novembre 2010 - 14:47 &bullet; SIGNALER UN ABUS - PERMALINK



*Grace a facebook, on risque d'atteindre l'immortalité in silico: lorsque l'IA sera capable de relire et comprendre tout ce qu'on a dit, on pourra nous modéliser assez finement apres 30 ans d'activite sur facebook*

*<http://intelligenceetfrustration.blogspot.com/2010/11/immortalite-passera-par-facebook.html>*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**ZIAD**

le 23 novembre 2010 - 20:42 &bullet; SIGNALER UN ABUS - PERMALINK



*Je me disais en lisant ce billet : qui peut bien être ce blogueur si savant et passionnant? Petit scroll pour remonter au nom et bien entendu, c'est Olivier Ertzscheid... quel talent! (et merci)*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**LAURA**

le 25 novembre 2010 - 17:04 &bullet; SIGNALER UN ABUS - PERMALINK



*Et là je regrette d'avoir effacé mon compte pour finalement en ouvrir un second quelques mois après...*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

**GOUPIL**

le 31 décembre 2010 - 15:56 &bullet; SIGNALER UN ABUS - PERMALINK



*rigolo de constater qu'"OWNI" est également présent sur facebook !...*

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

## 2 pings

Veille technologique du 18 novembre au 20 novembre le 20 novembre 2010 - 10:03

*[...] Facebook, la vie téléchargée , par @affordance – [...]*

Facebook, vos amis, vos données, votre vie | MediaNaranja le 23 novembre 2010 - 9:00

*[...] autre côté, Facebook vient de mettre à disposition une nouvelle fonctionnalité qui permet de télécharger l'ensemble des données nous appartenant que nous avons partagé [...]*