

Libroscope, plus que des
logiciels libres, des hommes
libres !

-- Réflexions --

Réflexions

À qui appartiennent vos données ?

Réponse peut-être intelligente
à une question apparemment
idiote

Raphaël Rousseau
mercredi 23 avril 2003

[1]

La réponse que j'attends, pour vous montrer le contraire naturellement, est « à moi ! ». Nous verrons dans cet article combien les contenus que renferment vos fichiers de données peuvent être liés à vos choix logiciels, et les dangers inhérents aux **formats** fermés.

Rappels et vulgarisation techniques

Tel que je le définissais sur [Wikipedia](#) [2] : *Un format de fichier consiste en la description (spécification) de la suite d'informations qui constituent un **fichier informatique**. Grâce à une telle description, il devient possible de définir un **programme informatique** capable de lire ou d'écrire de tels fichiers.*

C'est assez simple, non ? Les données sont codées sous forme numérique et stockées dans un fichier, qui permet d'y accéder globalement, des les manipuler, de les copier... Bien des formats reconnus par les logiciels sont des formats dits **binaires**, ce qui revient à dire qu'il ne suffit pas d'un simple **éditeur de fichiers texte** pour les ouvrir, mais d'un logiciel qui a été prévu pour gérer spécifiquement ce format.

Par conséquent, il revient à l'individu ou l'organisation qui est responsable du logiciel de décider si vous pouvez savoir comment sont organisées les informations à l'intérieur de vos propres fichiers.

Conséquences techniques

En présence d'un fichier, il faut pouvoir déterminer si on peut ou non en faire usage :

- un logiciel capable de gérer ce format est-il disponible ?
- les spécifications de ce format sont-elles disponibles ?

Avoir accès à ces spécifications permettrait à *quelqu'un* [3] de concevoir et réaliser un outil informatique capable de gérer le format en question, donc notre fichier.

Dans le cas d'une réponse négative à ces deux questions, deux réactions sont envisageables :

- abandonner son fichier, ce qui revient à gâcher la somme (parfois colossale) de travail qu'il a demandé pour être constitué ;
- ou passer par des chemins détournés [4].

Conséquences organisationnelles et stratégiques

Avez-vous songé que l'éditeur de vos logiciels favoris pourrait faire faillite et mettre la clef sous la porte en emportant dans sa tombe les spécifications des formats de données de vos fichiers ? En cas de perte totale de vos données sur votre disque dur, logiciels y compris, ne vous laissant que vos sauvegardes de données [5], mais pas nécessairement les logiciels idoines pour les manipuler ; allez-vous pouvoir éviter de les racheter ? Avez-vous pensé que votre éditeur de logiciel pourrait

changer de politique tarifaire et propose ses futures versions de votre logiciel à des tarifs prohibitifs ?
À travers vos données, ne seriez-vous pas très, voire complètement, dépendant des éditeurs de vos logiciels ?

Dans tous ces cas vous vous rendez bien compte que vos fichiers de données sont intimement liées aux logiciels qui les ont générés. Sans formats de fichiers ouverts, il y aura toujours une concurrence déloyale entre les éditeurs dominants [6] et tout producteur de logiciels alternatifs [7] : je parle de concurrence déloyale car un nouvel arrivant sur le marché pour un type donné de logiciels ne sera pas en mesure de rivaliser avec le leader du moment.

Même si vous avez les moyens

Admettons que vous soyez (employé par) une grosse entreprise, donc certainement grande utilisatrice de logiciels. Vous allez me rétorquer : *Nous ne misons que sur des fournisseurs pérennes, notre service achats est là pour ça. Certes nous perdons une part de notre indépendance en nous liant pieds et poings avec de grands éditeurs, mais il faut bien être dépendant de quelqu'un... Toujours est-il que votre propos ne nous concerne pas, ou pas de manière assez percutante !*

Très bien ! Si vous êtes de ceux qui n'ont pas de soucis pécuniers et pour qui l'achat massif de licences de logiciels est monnaie courante [8], lisez ce qui suit : vos politiques internes de gestion des logiciels sont-elles sûres ?

Certains logiciels comportent des fonctions de *sauvegarde rapide* qui mettent à jour le fichier sans en effacer effectivement le contenu préalablement présent dans le document. Au lieu de ça, les nouvelles informations ou les corrections apportées sont ajoutées en fin de fichier de manière assez sommaire, mais rapide, de sorte que l'examen du contenu du fichier par un simple [éditeur de texte](#) laisse voir ce que contenait votre fichier avant sa modification ! Eh oui, vous envoyez à l'entreprise **XX** un contrat lui garantissant l'exclusivité de tel ou tel marché, mais votre fichier peut vous trahir pour peu que votre interlocuteur prenne la peine de l'osculter de plus près : il y verra peut-être que ce même courrier a déjà été envoyé dans des termes similaires à l'entreprise **YY**. C'est juste que l'employé chargé de cette tâche a repris le fichier utilisé pour l'envoi destiné à **YY** et l'a modifié pour l'expédier à **XX**.

Cette même *fonctionnalité* [9] de sauvegarde rapide peut présenter un autre aspect tout aussi inquiétant. Lorsque vous réalisez une sauvegarde rapide, pour que cela soit vraiment rapide, le programme réalise un nombre d'opérations minimales, frôlant la négligence (et y sombrant parfois) : une partie de votre [mémoire vive](#) est directement placée dans votre fichier. Lors de cette opération, non seulement la partie de la mémoire contenant vos données se retrouve à l'endroit prévu, mais également des zones mémoires qui sont sensées être vierges. Or les programmes ne réinitialisent (généralement) pas la mémoire avant de l'utiliser ; on y retrouve donc des données laissées par le programme ayant précédemment utilisé cette zone. Si vous aviez utilisé un logiciel de comptabilité, ou de messagerie électronique par exemple, avant votre traitement de texte, il se pourrait que des zones exploitées par ce logiciel se trouvent recopiées à l'identique dans votre document, *alors qu'il n'a rien à voir* ! Ceci est réalisé de manière transparente pour l'utilisateur du traitement de texte, mais pas pour un éditeur de texte brut. Utilisez-vous des logiciels de chiffrement de vos messages électroniques, comme [GPG](#) ? Non ! Dommage, car vos mots de passes ou informations sensibles peuvent faire le tour du monde plusieurs fois par jours grâce à Internet, laissés à la merci des indiscrets qui sont aux points de circulation de l'information, comme les [routeurs d'Internet](#). [10]

Vite, une solution ! !

Nous voyons donc que l'accent principal est très souvent mis sur les logiciels et leurs fonctionnalités, par les éditeurs et même les testeurs sensés être impartiaux et experts dans ce domaine. Hélas, ils passent généralement sous silence les effets secondaires présentés ci-dessus qui résultent de chaque choix, à court comme à long terme. Pour ne pas risquer d'être pris à la gorge par n'importe quel éditeur, lilliputien ou gargantuesque [11], il reste le choix des formats ouverts qui garantissent :

- une [interopérabilité](#) entre logiciels, vous laissant l'option d'un choix sincère [12] pour les solutions logicielles,
- une sécurité plus grande, car vous aurez plusieurs moyens de savoir ce qu'il y a réellement dans vos fichiers,
- une garantie de pérennité dans vos investissements informatiques, car vous ne perdrez pas vos données en cas de changement de fournisseur de logiciels.

L'emploi ou non de formats de fichiers ouverts, ou de normes ouvertes en général, est le premier critère qui vous permette de juger de la confiance que vous pouvez mettre dans un fournisseur de solutions logicielles. Toute pratique visant à empêcher ou à compliquer l'exploitation de vos données sur des outils concurrents doit être une incitation à la méfiance. Naturellement, tous les [arguments les plus fallacieux](#) sont opposés par ces entreprises pour justifier votre captivité, mais *un homme averti en vaut deux* et maintenant vous ne pourrez pas dire qu'on ne vous avait pas prévenu...

Dans la même veine, l'article [1- Standards ouverts](#), [2 - Logiciels libres](#) apportera un éclairage entre logiciels libres et normes ouvertes.

Merci à Pierre Paccoud pour sa relecture attentive et ses remarques constructives qui m'ont permis de retoucher légèrement cet article, et de le rendre plus accessible et percutant.

[1] Le sous-titre s'inspire largement d'un livre de mon enfance intitulé *Réponses bêtes à des questions idiotes* (Editions Gallimard, 1983), dont le titre m'a toujours plu ;-)

[2] Pour en savoir plus sur Wikipedia, lire l'article [Wikipedia : une encyclopédie à part](#) .

[3] Ce quelqu'un est souvent appelé *développeur de logiciels*.

[4] Pour les chemins détournés, se pencher du côté de l'[ingénierie inverse](#) ; lorsque la Loi l'autorise

[5] Tout le monde fait des sauvegardes de ses données, placées en lieu sûr, non ?

[6] Ne m'accusez pas d'avoir pensé à Microsoft ! ;-)

[7] Je pense notamment aux logiciels libres, là.

[8] Passez-moi le jeu de mots...

À qui appartiennent vos données ?

[9] Une plaisanterie courante parmi les informaticiens est la question : *C'est un bug ou une fonctionnalité ?*. Ici, on peut légitimement se poser la question...

[10] Ces exemples m'ont été fournis lors d'une conférence de Roberto Di Cosmo lors d'une conférence à l'[EPFL](http://epfl.ch) le 12 juin 2001. Voir les transparents sur <http://elle.epfl.ch/archives/dicosm...>

[11] Gargantuesque par la taille, pas par l'appétit, naturellement...

[12] Référence à la campagne [Sincere Choice](#) initiée par Bruce Perens, un acteur très célèbre de la Communauté Internet, et particulièrement dans le milieu des logiciels libres.