

<https://dutailly.net/bureautique-quel-format-de-fichier>



Bureautique, quel format de fichier ?

- Culture numérique -

Date de mise en ligne : lundi 29 septembre 2014

Date de parution : 29 septembre 2014

Tutoriels et logiciels libres Licence Creative Commons Attribution - Partage

dans les Mêmes Conditions 4.0 International

Cette question des formats de fichiers est cruciale, pour le partage, mais surtout pour la pérennité des documents dont certains doivent pouvoir être gardés et lus dans le temps, voire retravaillés. C'est important pour tous les fichiers informatiques, quels qu'ils soient et pas uniquement les textes ou les tableurs [1]. Simplement ces documents sont très courants et aussi, c'est le domaine que je connais le mieux.

Un peu d'histoire pour commencer

Mis à part ma manie pontificatrice, un petit rappel de la courte histoire de la bureautique est essentielle pour comprendre l'importance du problème et ses enjeux.

Au début, c'est-à-dire vers la fin des années 1980 quand les ordinateurs ont commencé à envahir les bureaux et entamé leur entrée dans les foyers domestiques, il y avait une foultitude de systèmes informatiques (OS), qui ne se parlaient pas entre eux. Un document généré par un ordinateur d'un système donné ne pouvait être lu par un autre ordinateur d'un autre système. Pire, les disquettes, qui étaient le seul support mobile, n'étaient pas non plus compatibles entre elles, soit parce que les formats matériels étaient différents, soit parce qu'une disquette formatée pour un OS ne pouvait être lue par un ordinateur différent. La palme revenant à la firme Apple qui avait opté pour des disquettes 5 pouces 1/4 pour la série Apple II (et un OS particulier) et des disquettes 3 pouces et un autre OS pour la série Macintosh.



Différentes disquettes Existe en trois tailles. Source Wikimedia Commons : Floppy disk 2009 G1 - George Chernilevsky - Domaine public.

Évidemment, dans ces conditions, les logiciels avaient, eux aussi, des formats propriétaires et incompatibles entre eux.

Tant que les ordinateurs finalement, n'échangeaient pas entre eux, ou alors seulement en interne, cela ne portait pas à conséquence, sauf que... après une assez brève période d'effervescence, le secteur s'est resserré sur quatre OS principaux au début des années 1990 et les systèmes ont commencé à s'ouvrir les uns aux autres. Mais entre-temps, combien de documents restés dans les machines qu'il devenait impossible d'ouvrir ? Quelle quantité de travail perdu ? Quelles archives précieuses inaccessibles ?

Avec le début de l'informatique en réseau et des échanges de fichiers entre personnes diverses, il s'est avéré qu'il fallait trouver un format lisible par, disons, la majorité des logiciels de traitement de texte ou des tableurs.

Les premiers formats « universels » de texte

En matière de texte la hiérarchie était simple.

En bas de la pile, ce qui était recommandé en dernier lieu si tout le reste ne passait pas était le format **txt** ou texte brut. Il pouvait servir aussi pour les tableurs et les bases de données (et sert encore d'ailleurs) pour peu que les différents champs soient séparés par des séparateurs (virgule, point-virgule, etc.).

Avantages : très léger, il pouvait être lu de tout ce qui avait affaire à du texte.

Inconvénients : pas de mise en forme possible donc pas d'enrichissements typographiques (gras, italiques) et les caractères dits spéciaux, à savoir en gros tout ce qui ne figure ni dans la série des 26 lettres de l'alphabet ASCII, ni dans la liste des chiffres ou des ponctuations pouvaient être mal interprétés par le destinataire. Le résultat est peu agréable à l'oeil.

Ce format est encore utilisé en programmation, pour les bases de données et pour le contenu des sites internet, mais plus du tout (ou alors exceptionnellement) en traitement de texte.

C'est un format ouvert.

Le format **RTF** a été pendant un temps un bon compromis. Ce format a été développé par Microsoft. C'était le format de Wordpad, espèce d'intermédiaire entre un vrai traitement de texte et un vrai éditeur de texte.

Avantages : sans reprendre tous les éléments d'un document en traitement de texte complexe, notamment les styles, il garde les enrichissements typographiques (gras, italiques) et le rendu du texte est lisible (ce qui n'est pas vraiment le cas du txt).

Inconvénients : les fichiers générés sont assez lourds, c'est un format propriétaire dont Microsoft ne semble plus vouloir assurer l'évolution, c'est donc un format mort !

Le **.doc** : Microsoft ayant réussi à imposer son OS et sa suite bureautique à très grande échelle, le format ouvert odt étant, quant à lui encore balbutiant, c'est ce format propriétaire qui était conseillé pour des textes que l'on devait retravailler, voire tout court, les pdf étant encore relativement rares et il fallait le logiciel spécifique de la firme Adobe pour en générer. Cette hégémonie qui a commencé son essor à la fin du vingtième siècle devait céder le pas au début de cette décade 2010.

Avantages : format « universel de fait » dont on savait que le résultat serait identique d'un OS (Microsoft ou Apple) à l'autre pour peu que les machines fussent équipées de la même police de caractère.

Inconvénients : format propriétaire dont les logiciels n'existent pas pour Linux, remplacé en 2007 par le format docx et plus maintenu depuis cette année 2014. C'est donc un format mort à ne plus utiliser !

Formats « historiques » des feuilles de calcul

Concernant les feuilles de calcul, les choses sont très simples : au départ, si on devait ouvrir un tableau avec un autre logiciel, la seule solution envisageable était de le passer en texte séparé par des séparateurs ou **csv**. De cette façon il était possible de récupérer les données, pas le reste.

Là encore, on a eu un format « universel de fait » avec la montée en puissance du tableur Excel et la prédominance de la suite MSOffice qui, soit dit en passant, a laissé pas mal de cadavres derrière elle.

Inconvénients : le format **xls**, comme son confrère **doc** est un format mort car plus maintenu par le géant de Redmond.

Aujourd'hui quel format utiliser ?

À cette question du format à utiliser, il n'y a pas une réponse unique, mais il y a un principe. D'une façon générale, ce qui signifie qu'il peut y avoir des exceptions au cas par cas, **conseiller d'enregistrer systématiquement et de conserver les fichiers aux formats morts doc et xls est une faute** (et une ineptie). Voilà pour le principe.

Maintenant pour les réponses, il y a deux cas de figure.

1. Soit le fichier doit être **diffusé sans être retravaillé**, dans ce cas il doit être au format **pdf** (éventuellement au format epub pour du texte). Sans être libre, le format pdf s'est imposé comme norme de fait. Aujourd'hui, n'importe quelle suite bureautique actuelle est capable d'exporter un document au format pdf. Si ce n'est pas le cas de la vôtre, c'est que vous travaillez encore sur MSOffice 2003, qui n'est plus du tout maintenu par Microsoft depuis avril 2014. La solution changer de suite bureautique soit en faveur de la version plus récente du même éditeur, soit en version infonuagique (qui revient cher à la longue et vous oblige à être connecté en permanence), soit, enfin, pour des suites bureautiques libres qui n'impacteront pas votre budget.
2. Soit le fichier doit **être retravaillé ou est susceptible de l'être**, la meilleure solution, pour garantir sa pérennité, consiste à **l'enregistrer dans un format ouvert, à savoir odt pour le texte et ods pour les feuilles de calcul**. De cette façon, quelles que soient les évolutions, les spécifications du format étant accessibles, il sera toujours possible d'ouvrir le document dans un avenir plus ou moins lointain.

Aujourd'hui, la suite MSOffice pour Windows peut aussi enregistrer les documents dans des formats ouverts... pas celle pour OSX. Cela dit, évidemment, les suites bureautiques libres, elles, non seulement ouvrent les documents Microsoft et peuvent les travailler, mais elles sont disponibles sur tous les OS d'ordinateur et enregistrent par défaut les documents dans les formats ouverts.

Format ouvert ?

Les formats ouverts sont l'équivalent informatique du système métrique. Avant le système métrique, tous les systèmes de mesures étaient « propriétaires » et on ne pouvait en obtenir les spécifications qu'auprès du propriétaire, ce qui était d'ailleurs la porte ouverte à diverses tromperies et fraudes [2]. Le système métrique, étant, quant à lui, un système « libre » que n'importe qui peut recalculer à partir du rayon de la terre, qui peut se calculer aussi [3] et dont les spécifications sont diffusées largement, notamment à l'école pour la France.

À l'instar du système métrique, les spécifications des formats ouverts sont accessibles de tous, gratuitement, sans restriction et peuvent être utilisées par tous sans autre formalité. Ils garantissent ainsi non seulement la pérennité des documents mais aussi le fait qu'ils puissent être ouverts et travaillés quels que soient les systèmes (interopérabilité).

La notion est d'ailleurs très clairement définie par le droit français dans l'article 4 de la loi n°2004-575 du 21 juin 2004

pour la confiance dans l'économie numérique (la fameuse LCEN) :

On entend par standard ouvert tout protocole de communication, d'interconnexion ou d'échange et tout format de données interopérable et dont les spécifications techniques sont publiques et sans restriction d'accès ni de mise en oeuvre.

Les autres définitions légales sont très voisines, bien que rédigées différemment.

En conclusion

- Uniquement pour la diffusion : pdf
- Pour l'archivage en prévision d'un retravail éventuel : privilégier les formats ouverts
- Pour continuer à travailler sur le document : son format actuel sauf si c'est un .doc ou un .xls et qu'on n'utilise plus les vieilles suites MSOffice. Dans ce cas l'enregistrer au format docx ou xlsx quand on est sur Office et ods et odt dans les autres cas.

Pour aller plus loin :

- Les [recommandations de l'université de Strasbourg](http://documentationlogicielle.u-strasbg.fr/DUN/IPM/Formats/co/utilisationFormatsFichiers.html) [http://documentationlogicielle.u-strasbg.fr/DUN/IPM/Formats/co/utilisationFormatsFichiers.html] en matière de format.
- L'article de Wikipédia sur les [formats ouverts](https://fr.wikipedia.org/wiki/Format_ouvert) [https://fr.wikipedia.org/wiki/Format_ouvert].
- Le site [formats-ouverts](http://formats-ouverts.org/) [http://formats-ouverts.org/] qui traite du sujet à qui on souhaite, un peu en retard, un bon anniversaire, et une nouvelle décade fructueuse.
- La page du site [histoire-image](http://www.histoire-image.org/site/etude_comp/etude_comp_detail.php?i=306) [http://www.histoire-image.org/site/etude_comp/etude_comp_detail.php?i=306] sur le système métrique, soit dit en passant, si vous êtes à Paris et que vous avez du temps devant vous, je vous suggère d'aller faire un tour [au musée des Arts et Métiers](http://www.arts-et-metiers.net/) [http://www.arts-et-metiers.net/] et vous comprendrez dans les salles (oui il y en a plusieurs) consacrées à la métrologie et celles dédiées à l'histoire de l'informatique l'étendue du problème.
- Le site de [LibreOffice](http://fr.libreoffice.org/) [http://fr.libreoffice.org/] parce que je ne vais pas redessiner un gâteau (surtout aussi moche) et que c'est son quatrième anniversaire.

[1] Pour les présentations c'est à la fois plus simple car l'histoire est plus courte, et plus compliqué car un fichier de présentation est multimédia.

[2] Soit dit en passant, aujourd'hui encore le système de mesures étatsunien a dû faire l'objet de traité internationaux pour une définition uniforme...

[3] Bon, j'en serais incapable personnellement, mais c'est possible néanmoins.