

Tableurs : sommaire

Le logiciel tableur représente des données numériques et alphanumériques et permet en particulier de les traiter au moyen de fonctions. Les applications de gestion de tableaux sont fournies pour diverses plates-formes par des producteurs commerciaux et sont aussi disponibles sous forme de projets source ouverte. Chacune possède son propre format.

En principe, on peut tenir compte de trois aspects différents dans les considérations sur l'archivage de tableurs (documents provenant de tableurs):

- La fonctionnalité (aspect calcul). Pour assurer celle-ci, il est indispensable d'archiver en format original du tableur, c'est-à-dire en [XLS](#) , [ODF](#) ou [OOXML](#) .
- La nature du tableau. Celle-ci peut être prise en considération en traitant la feuille de calcul comme une base de données. On est alors renvoyé aux formats d'archivage des bases de données. En particulier, la possibilité d'exporter en XML est assurée par le format [SQLX](#) .
- L'aspect visuel. Celui-ci peut être reproduit par une conversion en [PDF/A](#) .

Les deux dernières possibilités ne sont de facto guère pertinentes. La conversion en PDF/A doit même être évaluée de manière critique parce qu'elle suppose un grand effort pour le contrôle de la qualité. En outre l'ensemble de la fonctionnalité est perdu (voir à ce sujet la discussion sur le format [PDF/A](#)).

Aucun format de fichier idéal du point de vue archivistique n'a pu s'imposer jusqu'ici pour les données de tableur. Une solution provisoire pour l'archivage consiste à miser sur la rétrocompatibilité des applications commerciales (en règle générale sur plusieurs versions). Par conséquent, il y a toutes les raisons de penser qu'un fichier en format courant pourra être lu sans problème pendant les dix années à venir par les versions suivantes du programme. Il est donc inutile de procéder à une migration. Dans le cas d'un fichier dans une version ancienne du format, il faut procéder à une migration dans le format actuel au moyen du logiciel correspondant (selon les investigations du «Digital Preservation Testbed» [p. 20], il est préférable de sauter chaque fois une ou deux versions). Dans le cas d'applications périmées en format propriétaire, il est indispensable d'effectuer une migration dans le format d'une application actuelle.

Ces considérations s'appliquent en particulier aux formats [XLS](#) et [OOXML](#) de Microsoft Excel qui dominent le marché. Bien que ces formats ne répondent pas aux exigences de l'archivage, il est plus judicieux d'archiver les fichiers MS Excel dans le format d'origine plutôt que d'effectuer une migration dans un format potentiellement plus approprié mais n'ayant pas fait ses preuves et dont la diffusion est faible. La lisibilité de 10 ans, atteinte grâce à la compatibilité descendante, accorde au service d'archivage un délai lui permettant de chercher un format d'archivage futur plus approprié.

Cette approche suit deux principes généraux guidant le choix d'un format, à savoir éviter les migrations superflues et garder ouvertes le plus d'options possibles.

Formats examinés

- [XLS](#)
- [ODF](#)
- [OOXML](#)
- [PDF/A](#)

Recommandation

A titre de solution provisoire, il faut archiver dans le format d'origine, en général [XLS](#) , [ODF](#) ou [OOXML](#) . Les alternatives comme archiver dans des formats d'archivage [Base de données](#) ou convertir en [PDF/A](#) ne peuvent être considérées comme bonnes pratiques et doivent être évaluées de manière critique.

Bibliographie

Digital Preservation Testbed: From digital volatility to digital permanence. Preserving spreadsheets
2003

↗ <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/volatility-permanence-spreadsh-en.pdf>

[Contact](#)
[A propos](#)
[Impressum](#)
[Événements](#)
[Newsletter](#)
[RSS](#)

ODF issu de tableur

Pour les caractéristiques techniques du format, consulter la description exhaustive de [ODF](#) dans la partie consacrée aux fichiers texte. On trouvera ici une brève évaluation d'ODF pour son utilisation comme format d'archivage pour [tableurs](#).

Evaluation

Ouverture du format : 4

Ce format est une norme OASIS depuis 2005 et il est disponible publiquement depuis 2006 comme ISO/IEC 26300.

Licence libre : 3

Ce format contient des brevets. Ceux-ci sont mis en circulation pour toutes les versions futures, tant que durera la collaboration avec Oracle.

Les polices utilisées ne font pas partie du domaine public, ce qui risque de poser des problèmes de droits lors de l'utilisation future.

Diffusion : 2

Ce format ne s'est pas vraiment imposé bien qu'occasionnellement il fasse ou qu'il ait fait l'objet d'une promotion active dans le domaine public (Allemagne, Belgique, Argentine, etc.).

Fonctionnalités : 4

OpenOffice permet de créer des documents ODF et peut convertir en ODF divers formats de traitement de texte courants et obsolètes. Préserver la présentation des documents ODF pose cependant problème. De même, les documents ODF originaux sont présentés sous diverses formes par les différents visualiseurs. Cela est dû au fait que les polices ainsi que les espaces de couleurs ne sont pas encapsulés dans les fichiers ODF. Les fichiers ODF peuvent contenir des macros, ce qui peut également empêcher un affichage fiable des documents. L'option de protection en écriture pose également des problèmes du point de vue archivistique.

Implémentation : 3

ODF est le format natif d'OpenOffice et de ses dérivés. Depuis 2009, Microsoft Office également permet de produire des documents en ODF.

Densité de mémorisation : 2

L'utilisation de la compression ZIP permet d'atteindre une densité de mémorisation relativement élevée. ODF n'approche toutefois pas la densité de mémorisation d'un format binaire comme [XLS](#).

Vérifiabilité : 3

La reconnaissance peut se faire grâce à l'extension ainsi qu'à la séquence d'octets correspondante. Il existe un validateur ODF.

Bonnes pratiques : 2

L'utilisation de ce format par les autorités et dans l'administration est encouragé par l'Etat dans divers pays. Le format est utilisé dans une faible proportion par les services d'archivage.

Perspectives : 2

En raison de la diffusion et de l'ouverture ainsi que du défaut d'alternative claire et établie, on estime que les perspectives sont relativement élevées.

Classe de formats : B

Ce format est intéressant pour l'archivage. Il existe depuis 2005.

Conclusion

ODF est un candidat possible pour l'archivage lorsque la fonctionnalité du tableur doit être conservée. Consulter la discussion au sujet de [XLS](#).

[Contact](#)
[A propos](#)
[Impressum](#)
[Événements](#)
[Newsletter](#)
[RSS](#)

OOXML issu de tableur

Pour les caractéristiques techniques du format, consulter la description exhaustive de [OOXML](#) dans la partie consacrée aux fichiers texte. On trouvera ici une brève évaluation d'OOXML pour son utilisation comme format d'archivage pour [tableurs](#) (XSLX).

Evaluation

Ouverture du format : 4

[OOXML](#) est disponible publiquement; c'est une norme Ecma International (Ecma-376) depuis 2006 ainsi qu'une norme ISO/IEC (ISO/IEC 29500:2012) publiée depuis 2012.

Licence libre : 3

Ce format contient des brevets. Ils ne sont chaque fois mis en circulation que pour la version actuelle.

Diffusion : 4

XLSX est de loin le format qui domine le marché pour les applications de tableurs.

Fonctionnalités : 4

La fonctionnalité de l'application de gestion de tableaux est assurée par le format d'origine.

Implémentation : 3

OOXML est le format d'origine de Microsoft Office, mais il peut également être généré par OpenOffice.

Densité de mémorisation : 3

L'utilisation de la compression ZIP permet d'atteindre une densité de mémorisation relativement élevée. OOXML n'approche toutefois pas la densité de mémorisation d'un format binaire comme [XLS](#).

Vérifiabilité : 2

La reconnaissance est possible uniquement au moyen de l'extension. Il n'existe aucun validateur spécial.

Bonnes pratiques : 3

Ce format n'est pris en charge dans l'archivage que dans une faible proportion. Etant donné qu'aucun format d'archivage ne se profile vraiment jusqu'ici dans le domaine des tableurs, de plus en plus d'institutions acceptent cependant OOXML au moins comme format d'archivage provisoire.

Perspectives : 3

En raison de la diffusion et de l'ouverture ainsi que du défaut d'alternative claire et établie, on estime que les perspectives sont relativement élevées.

Classe de formats : B

Format actuel

Conclusion

OOXML est un candidat possible pour l'archivage lorsque la fonctionnalité du tableur doit être conservée. Consulter aussi la discussion au sujet de [XLS](#).

PDF/A issu de tableur

Pour les caractéristiques techniques du format, consulter la description exhaustive de [PDF/A-1](#) et [PDF/A-2](#) dans la partie consacrée aux fichiers textuels. Comme le montre l'évaluation détaillée, il est déconseillé d'utiliser [PDF/A-3](#) pour l'archivage.

On trouvera ici une brève évaluation de PDF/A-1 et PDF/A-2 pour leur utilisation comme format d'archivage pour les [tableurs](#) .

Evaluation

Ouverture du format : 4

PDF/A est un standard ISO.

Licence libre : 3

En principe, ce critère est rempli pour PDF/A; il faut toutefois veiller à n'utiliser ni polices ni algorithmes de compression sous licence (algorithmes de chiffrement à partir de PDF/A-2).

Diffusion : 4

Les versions 1 et 2 de PDF/A sont entre-temps très répandues.

Fonctionnalités : 1

Seul l'aspect visuel des fichiers des tableurs peut être reproduit, la fonctionnalité est perdue. La conversion est souvent incomplète, par exemple parce qu'on a sélectionné une zone d'impression inappropriée ou lorsqu'il y a plusieurs feuilles de tableur dans un fichier.

Implémentation : 2

Il existe différents outils pour créer des documents PDF/A, mais la conversion est souvent incomplète (voir ci-dessus). Les fichiers PDF/A peuvent être lus par tous les lecteurs PDF.

Densité de mémorisation : 2

La densité de mémorisation est plus faible que celle des formats originaux.

Vérifiabilité : 3

Les fichiers PDF/A peuvent être reconnus au moyen de la séquence d'octets. Il existe de nombreux validateurs. On ne peut toutefois pas vérifier si toutes les informations du fichier original ont été converties.

Bonnes pratiques : 3

Ce format est déjà repris par les archives, mais il s'est avéré inapproprié pour les fichiers de [tableurs](#) en raison de la fonctionnalité limitée et des problèmes susceptibles d'arriver lors de la conversion.

Perspectives : 1

Les limites énoncées du PDF/A pour l'archivage de fichier de [tableurs](#) rendent irréaliste une utilisation répandue à cette fin.

Classe de formats : Ø

PDF/A est largement inapproprié pour l'archivage de fichiers de [tableurs](#)

Conclusion

PDF/A est un format qui a fait ses preuves pour l'archivage de [documents textuels](#) . Il ne convient cependant pas pour l'archivage de fichiers issus de [Tableurs-sommaire](#) à cause de la perte de toute la fonctionnalité et des problèmes susceptibles d'arriver lors de la conversion. A l'heure actuelle, il est impossible de contourner l'utilisation des formats originaux [XLS](#) , [ODF](#) ou [OOXML](#) .

[Contact](#)

[A propos](#)

[Impressum](#)

[Événements](#)

[Newsletter](#)

[RSS](#)

XLS

Informations générales

Titre	Microsoft Excel File Format; Binary Interchange File Format; (<i>Le format des fichiers Excel se nomme Binary Interchange File Format (BIFF). Mais comme Excel utilise l'extension de nom du fichier .xls, le format est mieux connu sous le nom XLS.</i>)		
Catégorie	Tableur		
Abréviation	XLS		
Extension de fichier	.xls		
Mime Type	application/msexcel; application/vnd.msexcel; application/excel		
Pronom PUID	fmt/55; fmt/56; fmt/57; fmt/58; fmt/59; fmt/60; fmt/61; fmt/62		
	<i>Version BIFF Version Excel Nom du logiciel (Windows)</i>		
	BIFF8	11.0	2003
	BIFF8	10.0	XP
	BIFF8	9.0	2000
Version	BIFF8	8.0	97
	BIFF5	7.0	95
	BIFF5	5.0	5.0
	BIFF4	4.0	4.0
	BIFF3	3.0	3.0
	BIFF2	2.0	2.x

Description

BIFF (XLS) était le format propriétaire de Microsoft Excel jusqu'à la version Excel 2003; il a ensuite été remplacé par [XLSX \(OOXML\)](#). Ses spécifications sont connues et ont été publiées par le projet OpenOffice.org; elles sont toutefois détenues par l'entreprise Microsoft. Ce format représentait la référence pour les tableurs jusqu'à ce qu'il soit remplacé par XLSX ; il jouissait d'une très grande diffusion sur différentes plates-formes. Les autres formats de tableur peuvent être convertis en XLS au moyen du logiciel approprié (convertisseur pour MS Excel). Bien que ce format soit périmé, il peut encore être lu et converti par les versions Excel les plus récentes.

Evaluation

Ouverture du format : 2

La spécification du cahier des charges du format a été obtenue par «reverse engineering», mais n'a pas été ouverte officiellement.

Licence libre : 2

XLS est un format propriétaire de Microsoft protégé par brevet. La spécification obtenue par «reverse engineering» est cependant utilisée «de facto».

Diffusion : 3

XLS est le prédécesseur de XLSX(OOXML) et il n'est en fait plus que répandu dans des fichiers anciens.

Fonctionnalités : 4

XLS assure la fonctionnalité de l'application en tant que format d'origine. Des convertisseurs pour d'autres formats de tableurs sont disponibles.

Implémentation : 3

XLS est principalement implémenté dans MS Office Excel. OpenOffice peut toutefois aussi lire les fichiers XLS.

Densité de mémorisation : 4

La densité de mémorisation est haute car il s'agit d'un format binaire.

Vérifiabilité : 2

Il est possible de reconnaître le format au moyen de l'extension de nom de fichier.

Bonnes pratiques : 3

Faute de format d'archive unique pour les tableurs, les documents Excel sont repris par les archives.

Perspectives : 2

En pratique, XLS est remplacé par XLSX(OOXML). En tant que format d'archive «de facto», il a cependant encore de l'avenir devant lui.

Classe de formats : A

XLS est le format pour tableurs qui domine dans les services administratifs.

Conclusion

Lorsque la fonctionnalité d'un document d'un tableur est considérée comme essentielle, il est même indispensable d'archiver dans un format d'origine XLS, ODF ou OOXML. Les documents qui sont déjà disponibles au format XLS ne doivent pas nécessairement être convertis au format Excel actuel, car la compatibilité descendante du logiciel garantit toujours leur lisibilité et permet ainsi un délai de migration vers un format mieux adapté à long terme. Malgré les problèmes d'archivage (format propriétaire ou spécification trop complexe), XLS peut être considéré comme un format d'archivage temporaire dans ce cas. Cependant, XLS ne devrait plus être utilisé comme format d'archive cible pour les migrations à partir de formats de tableurs obsolètes; à cette fin, il faudrait utiliser OOXML.

Références

OpenOffice.org's Documentation of the Microsoft Excel File Format. Revision 1.40. 2007

↗ <https://sc.openoffice.org/excelfileformat.pdf>

ECMA International, Standard ECMA-376, Office Open XML File Formats

↗ <https://www.ecma-international.org/publications/standards/Ecma-376.htm>

Bibliographie

Articles connexes

[OOXML issu de tableur](#)

