



Les formats de fichiers

1 Qu'est ce qu'un format de fichier ?

Chaque logiciel a sa manière propre d'encoder les informations.

Certaines manières d'encodage ne sont lisibles que par le logiciel lui-même.

D'autres sont lus par d'autres logiciels du même type, puis quelques formats de fichiers sont lus par tous les logiciels de la famille. Par exemple le format word Document (.doc) est produit et lu par Word mais aussi par d'autres logiciels de traitement de texte, mais pas par tous.

Alors que le format texte Mis en forme, ou RTF (.rtf) est lu et produit par tous les traitements de textes ou presque.

Les fichiers d'ordinateurs sont de deux types : ASCII ou binaires. Les premiers sont constitués de caractères qui respectent les codes ASCII. Ceux-ci définissent l'ensemble des caractères typographiques pouvant être reproduits à l'écran d'un ordinateur. En soi, les fichiers ASCII ne contiennent aucun élément de formatage ou de présentation (hormis les sauts de lignes). Ces fichiers ne sont pas limités à contenir du texte. Cependant, on pourra toujours les afficher avec un éditeur, même s'ils ne contiennent que de longues suites de caractères. Quant aux fichiers binaires, ils contiennent des informations codées directement en binaire (des 0 et des 1).



2 Quel format pour quel usage

Le problème du format de fichier ne se pose en fait que dans le cas où l'on désire envoyer ce fichier à quelqu'un. Le problème rencontré est alors souvent le fait que la personne ne possède pas le même logiciel et ne peut donc pas lire le document. Dans ce cas, il vaut mieux enregistrer le document dans un format partagé par beaucoup de logiciels.

Pour les fichiers texte, le format RTF est lisible par la plupart des logiciels de traitement de texte. Pour l'image, les formats Gif et Jpeg sont lisibles par les navigateurs internet. Pour le son et la vidéo, si on utilise un format utilisé sur internet, il y a en général moyen de télécharger un utilitaire qui permet de le lire.

4 L'importance de bien nommer les fichiers.

L'autre problème rencontré souvent quand on envoie ou reçoit un fichier est le nom qui lui a été donné. Il convient en effet de respecter certaines règles quand on donne un nom à un fichier. Les ordinateurs et les protocoles de transmission (mail, internet) se basent en partie sur le nom du fichier pour savoir comment le traiter. L'extension est alors très importante. L'extension, ce sont les quelques lettres, en général 3, qui se trouvent à la fin du nom, précédées d'un point. Un fichier au format RTF se présente comme : fichier.rtf



A titre d'exemple quelques extensions :

.doc	Word document	.mov	quicktime
.rtf	Texte mis en forme (Rich Texte File)	.mp3	MP3
.jpg	Jpeg	.png	PNG
.gif	Compuserve Gif	.swf	Shockwave Flash

Dans un envoi par mail, il est essentiel de bien mettre l'extension car le protocole de transmission peut abîmer le fichier s'il n'a pas cette indication sur la nature du format. Ce n'est pas systématique, mais cela arrive fréquemment. L'autre point à respecter est la règle suivante : dans le nom du fichier, ne jamais utiliser d'espace, de caractère de ponctuation excepté le tiret, le souligné et le point avant l'extension. Ne jamais utiliser de caractère accentué ni de caractère fantaisiste, type %, \$, *, etc.

Noms corrects :

monfichier.doc
mon_image.jpg
accordéon.mp3
film003.mov

fichier-copie.doc
animation_01_03.swf
exemple-test.rtf

Noms incorrects :

mon fichier.doc
mon_image
accordéon.mp3
film:003.mov

fichier.doc-copie
animation/01/03.swf
exemple&test.rtf

Erreur commise :

espace
pas d'extension
caractère accentué
ponctuation autre que le .
avant l'extension
ne termine pas avec l'extension
caractère spécial
caractère spécial

- L'utilisation des minuscules majuscules ne pose en principe pas de problème.
- Sachez seulement que les PC ne font pas la différence entre les minuscules et les majuscules, mais les Macintosh et les machines sous Unix ou Linux font cette différence.
- Il est donc plus prudent de toujours n'utiliser que des minuscules.

5 Les formats de fichiers pour documents textuels

Il s'agit sans doute de la catégorie de fichiers la moins bien normalisée.



.DOC (Types binaire ou ASCII)

Cette extension de fichier est utilisée pour représenter beaucoup de formats de textes. Le logiciel Microsoft Word, en version Windows, utilise l'extension .doc pour ses fichiers. Ces derniers sont assez répandus dans Internet.

Utiliser : Wordviewer (Windows) ; Logiciels de traitement de texte.



.HTML .HTM .SHTM .SHTML (Hypertext Markup Language)

(Type ASCII)

Utilisé spécifiquement pour la rédaction de page Web. Il s'agit d'une application de la norme SGML. (Les fichiers se terminant par .SHTM et .SHTML désignent des fichiers qui contiennent des instructions dynamiques de type SSI (Server Side Include). Ces instructions dynamiques pourraient être l'affichage de la date courante, le nombre de fois que la page a été accédée, ou même une information extraite d'une base de données.)

Utilisation : Navigateurs Web.



.PDF (Portable Document Format) - (Type binaire)

Très répandu pour la diffusion de documents dans Internet.

Utilisation : Acrobat Reader.

.PPT .PPS (PowerPoint Présentation, PowerPoint Slide Show)

(Type binaire)

Format du logiciel PowerPoint de Microsoft (suite Office), utilisé pour la production de diapositives électroniques généralement projetées sur un écran lors de présentations. Ce logiciel permet aussi de produire des « séances de projection » (slide shows) où les diapositives, qui contiennent souvent des images encapsulées, défilent automatiquement.

Utilisation : Microsoft PowerPoint Viewer.

.RTF (Rich Text Format) (Type ASCII)

Format de documents textuels normalisé favorisant l'échange entre différents logiciels et plateformes.

Utilisation : Logiciels de traitement de texte.



.QXD (QuarkXPress)

Extension de Quark, ce logiciel permet d'effectuer de la mise en pages avancée de documents destinés à l'impression (création de brochures, d'emballages, d'affiches...). Fonctions avancées de manipulation de textes et d'images, de gestion typographique.

Des fonctions d'édition HTML et XML



.indd (InDesign)

Les documents InDesign (mise en page professionnelle créé avec Adobe InDesign ; inclut des informations de formatage de page, du contenu de page, des fichiers liés et des styles ; utilisé pour formater des livres, des magazines, des journaux, des flyers, des brochures, etc.) peuvent être ouverts avec InCopy, le programme d'écriture professionnel Adobe et être importés dans QuarkXpress en utilisant le plugin ID2Q.

Sous Windows, les fichiers de document portent l'extension .indd, les fichiers de modèle l'extension .indt, les fichiers de bibliothèque l'extension .indl et les fichiers de livre l'extension .indb.

Compatibilité des formats :

*Ces deux logiciels reconnaissent la plupart des formats bitmap :
tiff, jpeg, bmp, eps, pict, gif, swf, psd, ai, pdf ...*

.TXT (Type ASCII)

Par convention, on donne cette extension aux fichiers de texte en simple format ASCII. Documents non formatés.

Utilisation : Navigateurs Web ; Logiciels de traitement de texte ; Éditeurs de texte variés (ex. Bloc-note pour Windows, SimpleText pour Mac).

.SGML .XML (Standard Generalized Markup Language, Extensible Markup Language) (Type ASCII)

SGML est une norme internationale pour la représentation de documents textuels en format électronique. Largement utilisé dans l'industrie de l'édition et les bibliothèques numériques pour la production et l'archivage de documents électroniques (périodiques, monographies, etc.). XML est une version « allégée » de SGML. Consulter le site The XML Cover Pages pour de plus amples informations

Utilisation : Navigateur Web : en général, les documents XML et SGML qui sont diffusés dans le Web sont convertis en HTML. Les versions récentes des navigateurs peuvent afficher directement les fichiers XML.

6

Les formats de fichiers pour images

BMP (BitMaP)

Le format BMP est le format par défaut du logiciel Windows. C'est un format matriciel. Les images ne sont pas compressées.

DXF

Le format DXF est un format vectoriel créé par la compagnie AutoDesk pour son logiciel de CAO AUTOCAD. Bien qu'étant un format très répandu dans le monde de la conception et du dessin assisté par ordinateur, le format DXF est très peu répandu en d'autres domaines.



.PSD (BitMaP)

Le format .PSD est le format d'enregistrement natif de Photoshop. C'est un format matriciel. Les images ne sont pas compressées.

.ai

Le format .ai est le format vectoriel d'enregistrement natif d'Illustrator.



.EPS (Encapsulated PostScript)

Un document en format EPS vectoriel est un fichier en langage PostScript décrivant le contenu d'une image ou d'une page complète (textes, filets, images importées). Contrairement aux fichiers PostScript (format .PS) qui ne peuvent être destinés qu'à l'impression, le fichier EPS peut être visualisé et importé dans bon nombre de logiciels de dessins, de PAO et de mise en pages. En fait, chaque fichier EPS contient une représentation matricielle en basse résolution en TIFF ou en PICT pour visualiser l'illustration à l'écran. Il faut noter qu'un fichier EPS est verrouillé. C'est donc dire qu'un fichier EPS exporté d'XPress ou d'InDesign ne pourra être modifié même lorsque ré-importé dans son logiciel d'origine. Seul un document Illustrator exporté en EPS vectoriel pourra être réouvert sur ce même logiciel et être modifié. Adobe Photoshop est également capable de modifier le contenu d'un fichier EPS vectoriel mais de ce fait, il le transforme en EPS matriciel.

Le format EPS matriciel n'est pas très différent du EPS vectoriel. En fait seules les données contenues dans le fichier sont différentes. Ainsi un logiciel de retouche de photos tel que Photoshop permet l'importation, la modification et l'exportation de fichiers en format EPS. À noter que le format EPS matriciel est sensible à la résolution de l'image.

Le fichier DCS (Desktop Color Séparation)

Il permet de proposer les images préséparées aux logiciels n'acceptant pas les images en EPS CMJN.

Les images à séparations multiples (5 ou 6 couleurs) CMJN + 1 ou 2 tons Pantone seront enregistrées en DCS 2 - paramètre EPS.

Le fichier PS (Post Script standard)

C'est un format de fichier que l'on peut transmettre directement aux unités de sortie ou aux serveurs d'impression, il est aussi utilisé pour transmettre des documents déjà mis en page vers des logiciels assurant des tâches telles que : l'imposition ou le trapping. Contrairement à l'EPS qui ne peut contenir qu'une page, un fichier PS peut contenir un document multipage, par contre il ne dispose pas de prévisualisation.

.GIF (Graphics Interchange Format) *(Type binaire)*



Format développé par CompuServe basé sur l'algorithme de compression LZW d'Unisys qui en détient les droits. Utilisé pour produire de petites images peu texturées, des icônes, des boutons dans les pages Web, etc. Permet la transparence et les images animées.



Le format GIF est un format qui a ouvert la voie à l'image sur le World Wide Web. Les images RVB ou CMJN en milliers de couleurs doivent d'abord être converties en 256 couleurs avant d'être exportées en format GIF. Le format GIF est un format qui à l'origine n'existait que dans le monde des IBM/compatibles. Toutefois, depuis quelques années, plusieurs logiciels Macintosh sont capables d'importer, de modifier et d'exporter des images en format GIF. Le format GIF ne peut toutefois pas être utilisé par les logiciels de mise en pages et de PAO.



Utilisation : Navigateurs Web

.JPEG (Joint Photographic Experts Group) *(Type binaire)*

Les images JPEG sont des images de 24 bits. C'est-à-dire qu'elles peuvent afficher un spectre de 16 millions de couleurs. C'est la meilleure qualité d'images disponible. Par contre, si la carte graphique de votre ordinateur est ajustée à 256 couleurs ou ne peut afficher plus de 256 couleurs, les images JPEG auront moins bonne mine que les images en format «GIF», qui sont pourtant des images en 8 bits (256 couleurs). Ce format accepte les images RGB et CYMK. Il est sans doute le mode compression le plus efficace qui soit, avec un bon compromis entre gain d'espace disque, temps de compression/décompression et qualité des images. Ainsi une image brute de 2Mo n'occupera après conversion en JPEG que 130 à 400 Ko selon la qualité d'image voulue. L'inconvénient des images JPEG, c'est qu'elles ne peuvent être importées

directement dans un logiciel de mise en page. Le format JPEG est un des formats les plus utilisés dans le monde du World Wide Web. Il peut être lu directement par les principaux outils de navigation pour le W3. Avantages et caractéristiques : Excellent ratio de compression, grande résolution tonale, utilisant un système 24 bits pour enregistrer la couleur (16.777.216 couleurs), excellent pour comprimer de gros fichiers images.

Utilisation : Navigateurs Web

PICT

Le format PICT est un format vectoriel interne au fonctionnement du Macintosh. C'est le format utilisé par le Presse-Papier du Macintosh. Il peut contenir des éléments graphiques ou des images numérisées. Ce format a quelques problèmes au niveau de la gestion de la séparation des couleurs, étape nécessaire à l'impression couleur. Pour des illustrations en noir et blanc ou en tons de gris, ce format peut satisfaire la majorité des besoins. Cependant, il est à proscrire dans le cas des impressions couleurs. Le format PICT est un format vectoriel d'où le gain appréciable en espace disque. C'est un format facile d'utilisation et reconnu par un grand nombre de logiciels autant sur environnement Macintosh que sur Windows. Il peut être utilisé en tant que format de travail, afin de faciliter les transferts entre logiciels. À ce titre, c'est le format de travail utilisé par Word pour Macintosh et son utilitaire MSGraph.

.TIFF .TIF (Tagged Image File Format) *(Type binaire)*

Le format TIFF, conçu à l'origine par la compagnie Aldus est un format matriciel. Conçu au départ pour n'accepter que les images en RGB, ce format permet de coder des images CMJN. Ainsi une image CMJN enregistrée en format TIFF peut être placée dans un logiciel de mise en pages et être envoyée à l'impression sans perte de qualité au niveau de l'image, compte tenu de sa résolution d'origine. Le format TIFF offre l'avantage d'occuper moins d'espace disque, grâce à son propre algorithme de compression appelé LZW.



Ainsi un fichier CMJN en format EPS occupe 25% moins d'espace après une conversion appropriée en format TIFF. Actuellement, c'est le choix à privilégier en PAO mais c'est aussi un choix à éviter pour le Web puisqu'aucun navigateur Web ne le lit directement.

Utilisé pour l'archivage d'images de qualité. Utilise un algorithme de compression sans perte. Fichiers de relativement grandes dimensions.



.PNG (Portable Network Graphics)

[Type binaire]. Appelé à remplacer les formats GIF et JPEG. Norme publique.

Utiliser : Windows : IrfanView. Mac : SimpleImage

WPG

Le format WPG est un format utilisé par les logiciels de la gamme de WordPerfect (WordPerfect, DrawPerfect, WP Presentations et autres) sous DOS, Windows ou Macintosh. C'est un format vectoriel qui donne un résultat acceptable lors de l'impression, mais qui doit surtout être utilisé en tant que format de travail. D'autant plus que ce n'est pas un format qui est reconnu par tous les logiciels.



PCX

Le format PCX est utilisé par le logiciel Paintbrush sous Windows. C'est un format matriciel.

7

Les formats de fichiers audio

.AIFF .AIF (Audio Interchange File Format) (Type binaire)

Relativement répandu dans Internet.

Utilisation : Netscape, Internet Explorer. Aussi Windows : Lecteur multimédia. Mac : lecteur QuickTime



.AU (Type binaire)

Développé pour Unix. Très répandu dans Internet.

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime

.MID (Music Instrument Digital Interface) (Type binaire)

Norme pour transmettre des informations musicales entre instruments électroniques et ordinateurs.

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime



.MP3 (Moving Pictures Expert Group) (Type binaire)

Format public pour la numérisation audio en haute-fidélité (qualité comparable au disque compact).

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime.



.RA .RAM (RealAudio) (Type binaire)

Format propriétaire pour la diffusion en direct de séquences sonores et vidéo par Internet.

Utilisation : RealAudio



.WAV (Waveform Audio File Format) (Type binaire)

Format utilisé par Windows pour stocker des séquences sonores.

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime

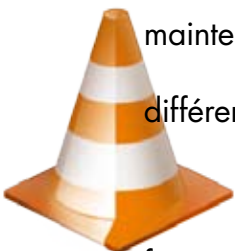
8

Les formats de fichiers vidéo

Il ne faut pas s'attendre à trouver des longs métrages en Dolby stéréo dans Internet (même si, de manière marginale, il est en fait possible d'en trouver). Les réseaux informatiques actuels ne permettent pas des débits de données suffisants. Néanmoins, la vidéo y est bien présente. Il existe deux méthodes de visualisation de séquences vidéo : la méthode en différé et la diffusion en direct. Avec la première méthode, il faut que tout le fichier ait été téléchargé sur son disque dur avant que l'on puisse le visionner. La seconde méthode, au contraire, permet la visualisation pendant le téléchargement. On comprendra que la qualité des séquences en direct sera médiocre. De plus, cette dernière nécessite une connexion très rapide à Internet, sinon les « projections » risquent d'être saccadées. Les séquences vidéo transmises en différé pourront être de meilleure qualité car le temps de téléchargement n'a pas à être synchronisé avec le temps de visualisation. (Selon le débit de la connexion Internet, il faudra peut-être une heure pour télécharger une séquence de quelques minutes.)

Concernant les logiciels de visualisation de vidéos, la plupart de ceux-ci prennent maintenant en charge les formats courants dans le Web.

Il n'est pas nécessaire d'avoir chacun des logiciels pour faire jouer des vidéos de ces différents formats. En général, un seul suffit.



.AVI (Audio Video Interleave : Video For Windows) (Type binaire)

Format vidéo générique pour Windows. Généralement visualisation en différé. Ce format peut faire appel à différentes techniques de compression de données grâce à l'utilisation de codec (codeur-décodeur) spécifiques à ces méthodes. (De ce fait, un fichier AVI particulier peut ne pas être lisible si le codec auquel il fait appel n'est pas présent sur son ordinateur.)

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime

.MOV .QT (QuickTime Movie) *(Type binaire)*

Développé par Apple. Commun sur Internet. Excellente qualité d'image. Généralement visualisation en différé.

Utilisation : QuickTime



.MPEG .MPG (Moving Pictures Expert Group) *(Type binaire)*

Norme publique. Très répandu dans Internet. Généralement visualisation en différé.

Utilisation : Windows : Lecteur multimédia. Mac : lecteur QuickTime

.RA (RealAudio) *(Type binaire)*

Format propriétaire pour la diffusion en direct de séquences sonores et vidéo par Internet.

Utilisation : RealAudio

.VIV (VivoActive) *(Type binaire)*

Format propriétaire pour la diffusion en direct de séquences vidéo par Internet. Qualité d'image médiocre mais fichiers très compacts.

Utilisation : VivoActive Player



.SWF (Shockwave Flash)

Le lecteur Flash, intégré dans la plupart des navigateurs permet de lire directement ce type de fichier pour animation. SWF Opener est un lecteur permettant de lire des vidéos au format Flash stockées sur le disque dur, par exemple dans les fichiers temporaires du navigateur.

9

Les formats de Compression, archivage et encodage

Ce type de fichier sont d'une très grande utilité pour les usagers d'Internet.

La fonction de compression de fichiers consiste à réduire leur taille en appliquant un algorithme qui enregistre l'information sous une forme plus compacte. La fonction d'archivage, quant à elle, sert à regrouper plusieurs fichiers dans un même fichier de destination. Les utilitaires d'archivage permettent même de regrouper des répertoires, des sous-répertoires et tout leur contenu sous un même fichier. Généralement, ces deux fonctions de compression et d'archivage se retrouvent simultanément dans la même application.

L'encodage utilisé, consiste à convertir des fichiers binaires en fichiers ASCII pour les rendre transmissibles par courrier électronique ou dans Usenet.



.ACE *(Type binaire)*

Format propriétaire efficace conçu pour archiver et compresser d'énormes fichiers. Il permet le morcellement en fragments, qui sont rassemblés lors de la décompression.

Utilisation : Windows : QuickZip (décompression seulement)



.BIN *(Type binaire)*

Encodage de fichiers pour Macintosh. Format 8 bits.

Utilisation : Mac : StuffIt Expander. Windows : StuffIt Expander.



.EXE *(Type binaire)*

Fichiers exécutables pour DOS ou Windows. Il s'agit de programmes proprement dit. On trouve dans Internet des fichiers de ce type qui sont en fait des archives compressées « auto-extractibles ».

Utilisation : Windows seulement : Exécuter le fichier. Décompression automatique des fichiers contenus dans l'archive.

.HQX (BinHex) *(Type binaire)*

Encodage ASCII de fichiers binaires. Très commun dans Internet pour les fichiers destinés au Macintosh.

Utilisation : Mac : StuffIt Expander. Windows : StuffIt Expander.

.GZ .GZIP (GNU ZIP) *(Type binaire)*

Compression de fichiers pour la plate-forme UNIX.

Utilisation : Mac : StuffIt Expander. Windows : QuickZip

.RAR (probablement Roshal Archive, d'après le nom du créateur) *(Type binaire)*

Format propriétaire efficace pour l'archivage et la compression de très gros fichiers. Il permet le morcellement en fragments. Ceux-ci sont rassemblés lors de la décompression.

Utilisation : Mac : StuffIt Expander (décompression seulement). Windows : QuickZip (décompression seulement)

.SIT .SEA (StuffIt archives, self-extracting archives) *(Type binaire)*

Archivage de fichiers pour Macintosh, avec compression. Format propriétaire de Alladin Systems. Les fichiers .SEA sont « auto-extractibles ».

Utilisation : Mac : StuffIt Expander. Windows : StuffIt Expander.

.TAR (Tape archive) *(Type binaire)*

Archivage de fichiers pour UNIX, sans compression. Note : dans Internet, les fichiers TAR sont souvent compressés de surcroît. Ils prennent ainsi une extension double : .TAR.Z. Dans ces cas, il faut d'abord décompresser le fichier avant de le désarchiver.

Utilisation : Mac : StuffIt Expander. Windows : QuickZip

.UUE (UUencoding) *(Type binaire)*

Encodage ASCII (7 bits) de fichiers binaires. Très commun pour le transfert de fichiers par courrier électronique et dans les forums Usenet.

Utilisation : Mac : StuffIt Expander. Windows : QuickZip. Note. Les clients de courrier électronique et Usenet font souvent le décodage automatiquement.

.Z (Compress) *(Type binaire)*

Compression de UNIX.

Utilisation : Mac : StuffIt Expander. Windows : QuickZip

.ZIP *(Type binaire)*

Format de compression et d'archivage très commun dans Internet, développé pour DOS et Windows.

Utilisation : Mac : StuffIt Expander. Windows : QuickZip



Glossaire :

Extension : nom donné aux lettres qui suivent le point et le nom d'un fichier.

Par exemple «doc» pour un fichier Word, ce qui donne «monfichier.doc»

Fichier : document numérique

Format propriétaire : format de fichier déposé par brevet. Les producteurs de logiciels qui souhaitent utiliser des formats de fichiers dont ils ne sont pas propriétaires (lecture et enregistrement) doivent acheter le droit de les utiliser. Ceci explique un petit peu le prix élevé de certains logiciels comme Photoshop qui achète beaucoup de brevets.

Lien hypertexte : une zone texte sensible qui appelle un autre fichier quand on clique dessus

Protocole : ensemble de conventions grâce auxquelles deux ordinateurs peuvent communiquer entre eux. Les protocoles qu'on rencontre le plus souvent sont : FTP (File Transfer Protocol), POP (Post Office Protocol), PPP (Point to Point Protocol), SMTP (Simple Mail Transfer Protocol)...