

D1.4 : Pérenniser ses données

L'enregistrement et la sauvegarde

Quand on travaille sur un document, celui-ci se trouve dans la mémoire vive de l'ordinateur. Pour conserver sa production, il faut l'**enregistrer** sous forme de fichier sur un support de stockage (le disque dur par exemple). Par sécurité, il est conseillé de dupliquer également cet enregistrement sur un autre support : c'est ce qu'on appelle la **sauvegarde**.

Ouvrir un fichier consiste à le charger dans la mémoire de l'ordinateur.

Enregistrer un fichier consiste à l'écrire sur un support de stockage.

Sauvegarder un fichier consiste à le dupliquer pour en faire une copie de sécurité.

Quelques conseils pour sauvegarder :

- réaliser des sauvegardes sur des supports externes (support amovible ou réseau) ;

Si vous sauvegardez votre fichier dans un autre dossier de votre disque dur, vous ne pourrez rien récupérer en cas de défaillance du disque.

Un moyen simple de faire une sauvegarde sans support amovible est de s'envoyer le fichier par courriel. Il sera stocké sur son serveur de messagerie et sera accessible de n'importe quel poste connecté à Internet.

- sauvegarder régulièrement (toutes les semaines ou tous les mois) l'ensemble des fichiers de données ;

Il est inutile de sauvegarder les applications, vous pourrez les réinstaller en cas de problème.

- sauvegarder quotidiennement les fichiers modifiés dans la journée.

Sur une clé USB ou votre ENT.

La gestion des versions

Il peut être utile de conserver les différentes versions d'un document ou les étapes successives de sa mise au point. Pour cela, il y a deux possibilités : produire autant de fichiers que de versions en les nommant judicieusement ou utiliser la fonction de l'application qui permet d'enregistrer plusieurs versions dans un même fichier.

Enregistrer un fichier par version

- On peut enregistrer un fichier pour chaque étape d'un projet ;

Exemples :

cours_c2i_version1.pdf

cours_c2i_version2.pdf

cours_c2i_version3.pdf

- On peut enregistrer un fichier pour les différentes variantes d'un document.

Exemples :

dupont 2011 relance mai.odt

lettre de relance avril.odt

lettre relance 15 novembre 2010.odt

lettre relance dupont 2009.odt

Les bonnes habitudes pour nommer les versions des fichiers

- Toujours nommer le fichier en respectant le même ordre des informations (par exemple sujet, date, auteur,...).
- Même s'il est possible de consulter la date d'un fichier dans ses propriétés, il est souvent utile de la voir apparaître dans son nom. Il est conseillé de mettre la date à l'envers de façon à ce que les fichiers soient listés par ordre chronologique.
- S'il existe plusieurs collaborateurs, il peut être utile d'indiquer le nom du dernier contributeur dans le nom du fichier.

Si on applique ces recommandations, on obtient :

relance_20091021_dupont_karine.odt

relance_20101115_dupont_nathalie.odt

relance_20110429_durand_nathalie.odt

relance_20110515_dupont_karine.odt

Cette représentation est plus cohérente et les fichiers s'affichent dans l'ordre chronologique !

Enregistrer les différentes versions dans le même fichier

- Les applications de bureautique permettent en général d'enregistrer les différentes étapes d'un document dans un même fichier (fonction de *versioning*).
- A tout moment, il est possible de revenir à une version antérieure.

L'archivage

Attention à la terminologie !

- Dans le langage courant, l'archivage consiste à classer et à ranger des documents dans le but de les conserver.
- En informatique, une archive est un fichier qui possède certaines propriétés.

Une **archive** est un fichier unique pouvant contenir une arborescence de dossiers et des fichiers.

Créer une archive est le moyen de rassembler un ensemble de fichiers en un seul.

Cela peut s'avérer utile pour :

- envoyer plusieurs fichiers en une seule pièce jointe ;
- organiser ses sauvegardes (par exemple, en créant une archive regroupant toutes les activités du mois).

Lorsqu'on télécharge une application (ensemble de fichiers) elle se présente souvent sous la forme d'une archive exécutable qu'on peut installer simplement.

En général, les logiciels permettant de créer des archives proposent en même temps de compresser les fichiers.

- Cette compression est réalisée sans perte d'informations : elle utilise des équations mathématiques complexes qui recherchent les éléments répétitifs d'un fichier et les remplacent par un code plus court.
- Le taux de compression dépend de la nature des informations à compresser.

Le taux de compression est important sur un format non compressé (comme Texte ou Bitmap) ; il est insignifiant sur un format déjà compressé (comme JPEG).

Un **logiciel de compression** permet de :

- compresser un ou plusieurs fichiers sous forme d'archive ;
- décompresser une archive pour récupérer le ou les dossiers et fichiers originaux.

Attention, une archive est souvent compressée mais ce n'est pas une obligation !

- Avantage : cela diminue la taille de l'archive.
- Inconvénient : il faut décompresser l'archive pour récupérer les fichiers.

Il est possible de créer une archive auto-extractible : cela permet à un interlocuteur qui n'a pas de logiciel adapté de décompresser l'archive par un simple double-clic.

Les unités de mesure

Pour traiter ou stocker des données, l'ordinateur ne sait manipuler que deux états possibles. C'est le principe du codage binaire. Le **bit** (*binary digit*) est l'élément de base du stockage de l'information : il peut prendre deux valeurs souvent symbolisées par 0 et 1. Les différentes informations (caractères, nombres, images,...) sont codées en binaire et la place qu'elles occupent dans la mémoire est mesurée en octets. Un **octet** est composé de 8 bits.

Attention, Byte est le terme anglais pour octet (1 Byte = 8 bits).

Combien de valeurs peut-on coder sur un octet ?

- Chaque bit peut prendre deux valeurs. Pour un octet, le nombre de combinaisons possibles est $2^8 = 256$.

*Quelques ordres de grandeur :
un octet permet de coder un caractère (lettre, chiffre, ponctuation, etc.),
un entier compris entre -127 et 128
ou un pixel parmi 256 couleurs.*

Les différentes unités de stockage :

- un kilooctet (ko) = 1 000 octets ;
- un mégaoctet (Mo) = 1 000 ko ;
- un gigaoctet (Go) = 1 000 Mo ;
- un téraoctet (To) = 1 000 Go.

*Quelques ordres de grandeur :
un morceau de musique de 3 minutes au format MP3 : 3 Mo
une photo au format JPEG de 5 millions de pixels : 2 Mo
un document au format texte de 40 pages sans illustration : 200 ko*

Abus de notation

En réalité, pour passer d'une unité à l'unité inférieure, on multiplie par 1024 (ce qui correspond à 2^{10}) au lieu de 1000.

Les unités correctes sont donc :

- 1 kibiocet (Kio) = 2^{10} octets = 1 024 octets ;
- 1 mébioctet (Mio) = 2^{20} octets = 1 024 Kio = 1 048 576 octets ;
- 1 gibiocet (Gio) = 2^{30} octets = 1 024 Mio = 1 073 741 824 octets ;
- 1 tébioctet (Tio) = 2^{40} octets = 1 024 Gio = 1 099 511 627 776 octets.

Les supports de stockage

Les supports amovibles

- La **clé USB** est un support de stockage facile à utiliser et à transporter. Elle se connecte directement sur le port USB d'un ordinateur. Elle peut également contenir des données multimédia (musique, photo ou vidéo) et être lue directement par certaines chaînes hi-fi, autoradios ou lecteurs vidéo.
- La **carte mémoire** est un support de stockage utilisé dans les appareils photos numériques, consoles de jeux, téléphones portables, lecteurs MP3, etc. Elle peut être exploitée par les ordinateurs qui possèdent généralement un lecteur pour lire ou écrire sur ces cartes.

Exemples de cartes mémoires : SD (Secure Digital), CompactFlash, MS (Memory Stick), MMC (MultiMedia Card), etc.

- Le **disque dur externe** existe sous différentes formes :
 - nomade sans alimentation : de la taille d'un agenda de poche, il est facilement transportable ;

Le port USB transporte un petit courant électrique qui suffit à alimenter certains dispositifs.
 - avec alimentation : idéal pour les sauvegardes du poste de travail, il faut le brancher sur le secteur.
- Les **CD** (Compact Disc), **DVD** (Digital Versatile Disc) et **BD** (Blu-ray Disc) sont des disques optiques. De nos jours, ils sont surtout utilisés pour acquérir des ressources (logiciel, musique et vidéo) car leur utilisation comme support de sauvegarde est contraignante : lenteur pour graver, nombre de copies limité, etc. On distingue :
 - CD-ROM (Read Only Memory), DVD-ROM ou BD-ROM : support qui ne peut être que lu (gravé à sa fabrication) ;
 - CD-R (Recordable), DVD-R ou BD-R : support qui peut être enregistré une et une seule fois ;
 - CD-RW (ReWritable), DVD-RW ou BD-RE : support sur lequel on peut lire et écrire de nombreuses fois.

Les supports en réseau

- Le **disque réseau partagé** est un disque accessible via un réseau local.

Vous pouvez également partager un simple dossier avec d'autres utilisateurs de votre réseau local.
- Le **serveur de fichiers** est localisé sur un ordinateur du réseau Internet ; on peut y accéder avec un logiciel client FTP ou avec un navigateur gérant le protocole FTP.

Certains sites (ENT, bureau virtuel, plateforme de travail collaboratif, etc.) permettent de stocker vos fichiers sur des serveurs distants par l'intermédiaire d'une interface web.

Obsolescence des formats et supports

- Les formats évoluent et peuvent devenir obsolètes.

Pensez à convertir vos anciens fichiers dans les formats actuels.
- Les supports de stockage ont une durée de vie limitée.

Pensez à recopier vos sauvegardes sur des supports récents.

L'évolution des supports en quelques valeurs : en 1981, la disquette 5 pouces 1/4 (13 cm) a une capacité de 360 ko ; dans les années 80, la disquette 3 pouces 1/2 (9 cm) a une capacité de 1,44 Mo ; dans les années 90, le CD enregistrable (12 cm) a une capacité de 700 Mo et dans les années 2000, la clé USB a une capacité de 2 Go à 32 Go.