

Information et données

1.2 Gérer des données

Source : C2i Domaine 1 – Modifié par Christophe Bulot

Nathalie DENOS

Maître de conférences en informatique
UPMF - Grenoble 2

Karine SILINI

Maître de conférences en informatique
ULCO

Licence Creative Commons :

Paternité - Pas d'Utilisation Commerciale
Partage des Conditions Initiales à l'Identique

Table des matières

Table des matières	2
Organiser un espace de travail complexe	3
L'environnement de travail	4
Le poste de travail.....	5
La configuration du poste de travail	6
Le réseau	7
La connexion au réseau	8
L'installation des applications.....	9
Le choix des applications.....	10
Les environnements numériques	11
Les espaces de stockage.....	12
L'organisation des fichiers	13
Tenir compte des enjeux de l'interopérabilité.....	14
Le format des fichiers	15
L'interopérabilité.....	16
Les formats de documents	17
Les formats d'images	18
Les autres formats.....	19
Pérenniser ses données	20
L'enregistrement et la sauvegarde.....	21
La gestion des versions	22
L'archivage.....	23
Les unités de mesure.....	24
Les supports de stockage	25

Organiser un espace de travail complexe

L'environnement de travail

Avec le développement des technologies de l'information et de la communication (TIC), la plupart des configurations matérielles sont désormais connectées à Internet ou à un réseau local, ce qui permet notamment de **délocaliser** (installer sur des serveurs distants) un certain nombre de ressources et d'applications.

L'**environnement de travail** d'un usager se compose :

- d'un poste de travail avec ses périphériques : souris, clavier, imprimante, casque, etc. ;
- d'un ensemble d'applications, installées sur le poste de travail ou sur des serveurs distants ;
- d'espaces de stockage, locaux ou distants.

Dans son quotidien, l'usager passe d'un environnement local à un environnement distant (et vice versa) très facilement et souvent sans s'en apercevoir.

Quels peuvent être les services délocalisés ?

- Des logiciels de communication et de collaboration (messagerie, forum, agenda partagé, etc.).
- Des logiciels de bureautique en ligne (traitement de texte, tableur, etc.).
- Des espaces de stockage privés ou partagés.
- Des applications permettant de rester en contact et d'échanger avec des amis ou relations.

Quels sont les avantages de délocaliser les services ?

- L'accès aux services peut se faire de n'importe quel poste de travail connecté à Internet (ordinateur, téléphone, tablette, etc.).
- La maintenance des applications et la sauvegarde des données sont prises en charge par un prestataire.

« L'**informatique en nuage** ou *cloud computing* est une forme particulière de gérance de l'informatique, dans laquelle l'emplacement et le fonctionnement du nuage ne sont pas portés à la connaissance des clients ». Note extraite de la fiche « informatique en nuage » sur [FranceTerme](#) (consulté le 10 août 2011)

Le poste de travail

On peut différencier trois types d'ordinateurs selon leur usage.

- L'ordinateur du domicile : en général, c'est un ordinateur fixe équipé d'un certain nombre périphériques (moniteur, imprimante, scanner, disque,...). Il est souvent utilisé par plusieurs membres de la famille.
- L'ordinateur portable : c'est un ordinateur personnel que l'on emporte avec soi. Il intègre les périphériques standards (pavé souris, écran, clavier, caméra, micro, enceinte,...) mais est en général moins performant et moins confortable à utiliser que le fixe.
- L'ordinateur du travail : il s'agit souvent d'un poste fixe connecté au réseau de l'organisation.

Quelles sont les principales caractéristiques d'un ordinateur ?

- La puissance du **processeur** : elle détermine la rapidité de traitement.
- La capacité de la **mémoire vive** (RAM) : c'est la mémoire qui est utilisée pour l'exécution des programmes en cours.

C'est une mémoire volatile : elle s'efface quand on éteint l'ordinateur.

Plus la mémoire vive est importante, plus l'ordinateur est rapide.

- La capacité de **stockage**.
- Le **système d'exploitation** : c'est un ensemble de programmes assurant la liaison entre les ressources matérielles, l'utilisateur et les applications. Les systèmes d'exploitation les plus répandus sont :
 - Windows : système d'exploitation de Microsoft ; installé par défaut sur pratiquement tous les ordinateurs personnels (sauf ceux de la marque Apple), il détient actuellement le quasi monopole du marché grand public ;
 - Linux : système d'exploitation distribué selon les règles du logiciel libre ; il en existe différentes distributions (Ubuntu, Mandriva, Debian, Red hat , etc.) ;
 - Mac OS : système d'exploitation des ordinateurs de type Macintosh d' Apple.
- La **connectique**
 - Le port USB (Universal Serial Bus) : port permettant de connecter « à chaud » des périphériques à un ordinateur, c'est-à-dire sans avoir à redémarrer l'ordinateur et avec une procédure d'installation minimale.
 - Le port Ethernet (RJ45) : port permettant de connecter l'ordinateur à un réseau local.
 - Le port VGA (Video Graphics Array) : port permettant de connecter l'ordinateur à un écran en analogique.
 - Le port HDMI (High Definition Multimedia Interface) : port permettant de connecter l'ordinateur à un écran en numérique haute définition.
- Les **connexions sans fil**
 - Le Wi-Fi : technologie de réseau local sans fil à haut débit ; il est surtout utilisé pour connecter sans fil un ordinateur à Internet via une borne Wi-Fi.
 - Le Bluetooth : technologie radio courte distance ; il est surtout utilisé pour faire communiquer entre eux des appareils situés à proximité les uns des autres sans qu'un câble soit nécessaire.

Pour synchroniser le carnet d'adresses de son téléphone avec celui stocké sur son ordinateur, on peut utiliser une connexion Bluetooth.

La configuration du poste de travail

A la mise sous tension d'un ordinateur, les deux phases suivantes s'enchaînent :

- lancement du BIOS (Basic Input Output System), petit programme contenu sur la carte mère de l'ordinateur qui vérifie le bon fonctionnement du matériel ;
- chargement du système d'exploitation.

Ensuite, deux cas de figure peuvent se présenter :

- soit l'ordinateur gère plusieurs utilisateurs et demande une identification : après l'authentification, le système d'exploitation ouvre une session spécifique et donne l'accès aux ressources en fonction du profil de l'utilisateur (administrateur, usager, invité, ...);

La personne qui configure l'ordinateur familial crée un compte pour chacun avec des droits limités pour les plus jeunes.

- soit l'ordinateur est utilisé par une seule personne qui est alors l'administrateur du poste.

La configuration par défaut

Quand on acquiert un ordinateur, seul le système d'exploitation et quelques logiciels utilitaires font partie de la configuration de base.

- Le **panneau de configuration** (menu des préférences du système) permet de personnaliser l'affichage (résolution, couleur, thème, image de fond, etc.), gérer les comptes utilisateurs et paramétrer les périphériques (imprimante, caméra, micro et casque, etc.).

La résolution de l'écran est le nombre de pixels affichés : plus elle est importante, plus de choses sont affichées ... au détriment de la lisibilité.

- Le **gestionnaire de fichiers** permet de manipuler les fichiers et dossiers.
 - On peut paramétrer la présentation des fichiers : le modèle d'affichage (miniatures, listes, détails, etc.), l'ordre d'affichage, les informations à afficher (fichiers cachés, extensions, etc.).
 - On peut consulter ou modifier certaines propriétés des fichiers : les droits d'accès (en lecture seule, modifications autorisées, etc.), l'application associée par défaut à un type de fichier, etc.

Le gestionnaire de fichiers peut avoir des noms différents selon les systèmes : l'explorateur pour Windows, le navigateur de fichiers pour Linux ou le finder sous Mac

L'installation d'un périphérique

L'installation d'un nouveau périphérique peut nécessiter un **pilote informatique** ou *driver* : c'est un programme qui permet au système d'exploitation de gérer le périphérique. Il peut être fourni sur un CD d'installation livré avec le périphérique, se télécharger automatiquement sur Internet dans le cas d'un périphérique *Plug and Play*, ou être déjà présent dans la configuration par défaut de l'ordinateur.

Le réseau

Un **réseau informatique** est un ensemble d'équipements reliés entre eux pour échanger des informations.

Un **réseau local** est un réseau informatique se limitant à une pièce ou un bâtiment. Il est souvent composé de plusieurs ordinateurs ou périphériques reliés entre eux. Un **réseau étendu** est un réseau informatique couvrant une grande zone géographique qui peut s'étendre à la planète toute entière.

Internet est un réseau informatique mondial. Il résulte de l'interconnexion d'une multitude de réseaux informatiques à travers la planète.

Aucun élément d'Internet ne connaît le réseau dans son ensemble. La communication sur Internet est possible en utilisant un protocole commun qui permet l'acheminement des données de proche en proche.

Chaque ordinateur connecté à Internet dans le monde est identifié par une **adresse IP** (*Internet Protocol*) composée de 4 séries de chiffres séparés de points. Un **fournisseur d'accès à Internet** (FAI) est un prestataire de services qui met à disposition des adresses IP (temporaires ou fixes) pour pouvoir se connecter au réseau Internet.

Si vous êtes connecté à Internet à partir d'un réseau local (université, réseau domestique de type « box », etc.), vous partagez une adresse IP extérieure commune pour accéder à Internet et vous êtes identifié par une adresse IP locale attribuée par le serveur mandataire ou proxy du réseau local.

Pour pouvoir communiquer entre eux, les différents éléments du réseau vont utiliser des **protocoles de communication**.

Un **protocole réseau** définit de façon formelle et interopérable la manière dont les informations sont échangées.

On distingue le **protocole TCP/IP** (*Transmission Control Protocol/Internet Protocol*) qui est le protocole de transport des données d'un ordinateur à l'autre sur Internet et les **protocoles d'applications** qui déterminent ce qu'il faut envoyer sur le réseau et comment exploiter les données reçues.

Les **protocoles d'applications** sont :

- HTTP (*HyperText Transfer Protocol*) : protocole de transfert hypertexte pour naviguer sur le web (HTTPS pour la version sécurisée) ;
- FTP (*File Transfer Protocol*) : protocole destiné au transfert de fichiers informatiques ;
- SMTP (*Simple Mail Transfer Protocol*) : protocole utilisé pour transférer le courrier électronique vers les serveurs de messagerie ;
- POP3 (*Post Office Protocol version 3*) : protocole utilisé pour récupérer le courrier électronique d'un serveur de messagerie ;
- IMAP (*Internet Message Access Protocol*) : protocole utilisé pour consulter son courrier électronique sur un serveur de messagerie.

Attention, ne pas confondre Internet et ses services ! Le web ou le courriel sont des services qui utilisent le réseau Internet.

Un serveur web est un ordinateur qui contient des pages web et les met à la disposition du réseau Internet. Ces pages sont généralement reliées entre elles par des hyperliens. Le **web** (ou la toile) est l'ensemble des hyperliens qui relient les pages web entre elles.

La connexion au réseau

Pour connecter un ordinateur à un réseau, il y a plusieurs possibilités :

- la **connexion filaire** par le port Ethernet (RJ45) ;
- la **connexion sans fil** de technologie Wi-Fi ; pour cela, il faut que l'ordinateur dispose d'une carte Wi-Fi (intégrée ou externe) et se trouve à proximité d'une borne Wi-Fi ;

Pour des raisons d'économie d'énergie, la carte Wi-Fi des ordinateurs portables peut être désactivée. Il suffit de l'activer au moment de s'en servir, soit par une combinaison de touches, soit par un interrupteur.

- la **connexion par une clé 3G+** équipée d'une carte SIM via le réseau de téléphonie mobile.

La connexion en Wi-Fi

Si la carte Wi-Fi est activée, les réseaux sans fil disponibles s'affichent et il suffit d'en choisir un. On distingue :

- les réseaux sécurisés pour lesquels une authentification est requise ;

Par exemple, la connexion au Wi-Fi d'une « box » est sécurisée par une clé WEP ou WPA.

- les réseaux non sécurisés.

Lors d'une telle connexion, le navigateur peut être automatiquement redirigé vers un portail captif dont la page d'accueil demande une identification, voire un paiement. C'est le cas des bornes Wi-Fi que certains fournisseurs d'accès mettent à disposition de leurs abonnés dans les lieux publics.

Pour établir une connexion avec une borne Wi-Fi, il est conseillé de paramétrer l'ordinateur en mode itinérant, ce qui configure la connexion de façon dynamique grâce au protocole **DHCP** (Dynamic Host Configuration Protocol).

Est-on anonyme quand on se connecte via une borne Wi-Fi sans être identifié ? Non, la carte réseau de tout ordinateur est identifiée par une adresse MAC unique ...

L'installation des applications

En plus du système d'exploitation et des logiciels utilitaires fournis, il est utile d'installer un certain nombre d'applications ou logiciels sur un ordinateur.

Un **logiciel** est un ensemble de fichiers permettant d'exécuter un programme informatique.

Parmi les logiciels, on distingue :

- **les applications** : logiciels destinés aux utilisateurs comme le traitement de texte, le navigateur, etc ;
- **les logiciels systèmes** : logiciels proches de la machine qui permettent aux applications de communiquer avec le matériel.

Le système d'exploitation est un logiciel système de base. Chaque application est développée pour fonctionner avec un système d'exploitation spécifique.

L'installation d'une application s'effectue :

- soit à partir d'un CD d'installation en suivant la procédure décrite ;
- soit à partir d'un téléchargement sur le web : en général, un fichier exécutable est téléchargé ; son lancement installe l'application (création d'un dossier, création de raccourci, etc.) ; à la première utilisation de l'application, certaines questions permettent de personnaliser l'installation (identité de l'utilisateur, langue par défaut, etc.).

On ne peut installer une application que si on détient les droits d'administrateur de l'ordinateur.

A l'exécution d'un fichier téléchargé, une confirmation peut être demandée car les fichiers exécutables peuvent contenir des virus. Une fois l'application installée, on peut supprimer le fichier exécutable téléchargé.

Lors de l'installation d'un logiciel propriétaire, une clé de licence ou clé d'enregistrement est demandée ; elle est fournie lors de l'acquisition du logiciel.

La mise à jour des applications est souvent proposée automatiquement.

A partir du moment où l'ordinateur est connecté à Internet, la mise à jour la plupart des applications installées est proposée si une nouvelle version est disponible.

Régulièrement, des « erreurs de programmation » ou bugs et des « failles de sécurité » sont mises en évidence. Les mécanismes de mise à jour du système permettent de télécharger et d'installer les correctifs ou compléments.

Pour supprimer une application, il faut utiliser la procédure de désinstallation accessible :

- soit par le menu de l'application,
- soit par le gestionnaire de programme,

et qui permet d'enlever « proprement » l'application et tous les fichiers associés (menu, raccourci, mention dans la base de registres...).

Attention, il ne faut surtout pas supprimer directement le dossier de l'application, au risque de provoquer des comportements inattendus du système.

L'icône qui figure parfois sur le bureau pour lancer l'application n'est qu'un raccourci vers le logiciel ; le supprimer ne désinstalle pas le logiciel.

Le choix des applications

Il existe un certain nombre de logiciels qu'il est conseillé d'avoir sur son ordinateur.

Une **suite bureautique** est une application regroupant un ensemble de logiciels visant à automatiser les activités de bureau : traitement de texte, tableur, logiciel de présentation, gestionnaire de bases de données, logiciel de dessin, etc.

Exemples : Open Office (logiciel libre et gratuit) ou Microsoft Office (logiciel propriétaire).

Avant d'utiliser un logiciel de bureautique, vérifiez les paramètres : les données d'identité pour identifier la personne ayant créé ou modifié le document, les paramètres linguistiques pour définir la langue et les options de correction automatique, etc.

Un **navigateur Web** est un logiciel dont la fonction principale est de consulter les ressources du Web.

Exemples : Mozilla Firefox, Google Chrome (logiciels libres et gratuits) ou Internet Explorer (logiciel propriétaire).

Un **client de messagerie** est un logiciel permettant de lire et d'envoyer des courriels.

Exemples : Mozilla Thunderbird (logiciel libre et gratuit) ou Microsoft Outlook (logiciel propriétaire).

Pour configurer un client de messagerie, il suffit d'indiquer dans les paramètres le nom du serveur de courrier entrant, le nom du serveur de courrier sortant et l'adresse électronique.

Un **logiciel de traitement d'image** est un logiciel offrant la possibilité de modifier des images numériques. Il permet en général d'éditer une image, de changer son format, son taux de compression, ses dimensions, etc.

Exemples : Gimp (logiciel libre et gratuit), XnView (logiciel gratuit pour une utilisation non commerciale).

Un **lecteur PDF** est un logiciel qui permet de visualiser et d'imprimer un fichier au format PDF (Portable Document Format).

Exemple : Adobe Reader (logiciel propriétaire et gratuit).

Pour créer un PDF, vous pouvez utiliser le logiciel libre et gratuit « PDF creator » qui fonctionne comme une imprimante virtuelle générant des fichiers PDF.

Un **antivirus** est un logiciel conçu pour protéger les ordinateurs des logiciels malveillants.

Exemples : Avast! (gratuit pour une utilisation non commerciale), Norton Internet Security ou BitDefender (logiciels propriétaires).

Un **logiciel de compression** permet de compresser un ou plusieurs fichiers sous forme d'archive. Il permet également de décompresser une archive pour récupérer le ou les fichiers originaux.

Exemples : 7-Zip (logiciel libre et gratuit), WinZip, WinRAR (partagiciels), PowerArchiver (logiciel propriétaire).

Les environnements numériques

Grâce à Internet, l'utilisateur évolue dans des environnements numériques en ligne divers et variés.

La plupart d'entre eux proposent des services personnalisés et individualisés à partir du moment où l'utilisateur est identifié.

Parmi ces environnements numériques, on peut distinguer les environnements plus spécifiques à l'activité professionnelle :

- l'**environnement** ou **espace numérique de travail** (ENT) : dispositif global fournissant à un usager un point d'accès à travers les réseaux à l'ensemble des ressources et services numériques en rapport avec son activité ;

L'ENT n'offre en lui-même qu'un nombre restreint de services de base. Son rôle principal est d'intégrer les services et de les présenter aux utilisateurs de manière cohérente, unifiée, personnalisée et personnalisable.

- la **plateforme de travail collaboratif** : site regroupant des outils pour travailler à plusieurs à distance ;
- la **plateforme pédagogique**, ou **d'enseignement à distance** ou **e-learning** : site permettant l'apprentissage à distance : cours en ligne, activités pédagogiques, outils de communication entre enseignant, tuteur et apprenant, etc.

et les autres :

- le **réseau social** qui permet de rester en contact et d'échanger avec ses relations ;
- le **site collaboratif** ou **de partage** : forum, site de partage de photos, encyclopédie collaborative, etc. ;

Soyez vigilant ! Quand vous publiez un fichier dans ces environnements numériques en ligne, demandez-vous toujours qui pourra y avoir accès.

Autres définitions proches

Un **bureau virtuel** est un site web offrant des services d'organisation, de communication, de collaboration et de stockage de documents.

Un **intranet** est un réseau informatique à l'usage exclusif d'un organisme utilisant les protocoles et techniques d'Internet ; si ce réseau est ouvert à certains usagers extérieurs, on parle d'**extranet** ;

Les espaces de stockage

Pour pouvoir être conservée, toute information doit être stockée sous forme de fichier sur un support physique.

La mémoire vive de l'ordinateur ne permet pas de conserver l'information !

Parmi ces supports de stockage, on distingue :

- le **disque dur** de l'ordinateur ;
- les **supports amovibles** (clé USB, carte mémoire, disque dur externe, CD/DVD, etc.) : l'accès aux fichiers est possible à partir de n'importe quel poste de travail auquel le support est connecté ;
- les **espaces en ligne** (espace privé/partagé d'un ENT, espace de publication chez un hébergeur, serveur de fichiers distant, etc.) : l'accès aux fichiers est possible à partir de tout matériel connecté à Internet.

Toujours faire une copie des fichiers importants de son disque dur sur un support amovible ou sur un espace en ligne (ou les deux) : en cas de défaillance matérielle du disque dur, les fichiers seront récupérables.

L'organisation des fichiers

Un **fichier** est une suite d'informations stockées sur un support physique sous forme de blocs de données binaires.
Un **dossier** contient des fichiers ou d'autres dossiers.

En y regardant de plus près au niveau du stockage en mémoire, un dossier est un fichier particulier qui joue le rôle de conteneur de fichiers.

En imbriquant des dossiers, on peut créer une **arborescence** qui permet de ranger logiquement ses fichiers dans des dossiers et sous-dossiers.

Pour créer une arborescence de dossiers, il suffit de se placer dans le dossier « parent » et de créer un nouveau dossier.

Le **chemin d'accès** d'un fichier est la liste des dossiers à parcourir pour atteindre ce fichier.

Si cette liste commence à la racine (point d'entrée) du support, on parle de **chemin absolu** sinon c'est un **chemin relatif**.

La **désignation d'un fichier** est unique par la donnée de son chemin d'accès et de son nom.

Selon les systèmes d'exploitation, la racine et les séparateurs de dossiers diffèrent.

Exemples :

Sous windows : C:\Alex\C2i\Exercices\D1\Fiche.odt est la désignation exacte du fichier.

\Alex\C2i\Exercices\D1 est le chemin absolu car il part de la racine du support (C:)

..\C2i\Exercices\D1 est un chemin relatif car il dépend du dossier où l'on se situe ; .. signifie qu'il faut remonter au dossier parent ;

Sous linux : /home/Alex/C2i/Exercices/D1/Fiche.odt est la désignation exacte du fichier.

L'**extension** est le suffixe du nom de fichier. Il donne une information sur le type du fichier.

L'extension du fichier Photo.jpg est jpg, et permet de dire que c'est une image.

Avez-vous déjà vu deux fichiers de même nom dans un même dossier ? Si les extensions sont masquées par le gestionnaire de fichiers, les fichiers Photo.png et Photo.jpg apparaissent tous les deux sous le nom Photo.

Un **raccourci** ou alias ou lien symbolique est un lien vers un fichier ou un programme. L'icône d'un raccourci placé sur le bureau est souvent repérable par la présence d'une flèche.

La suppression d'un raccourci ne supprime pas le fichier vers lequel il pointe !

Les **libellés** ou **tags** permettent d'organiser des fichiers de façon plus souple que les dossiers, car on peut associer plusieurs libellés à un même fichier alors qu'on ne peut le mettre que dans un seul dossier.

Ce mode d'organisation est très répandu pour les fichiers en ligne, les signets ou photos partagés. Il prend de l'ampleur dans l'organisation des courriels et des documents dans les espaces de stockage partagés.

Tenir compte des enjeux de l'interopérabilité

Le format des fichiers

Le **format d'un fichier** décrit la façon dont l'information est codée dans le fichier.

Il ne faut pas confondre le **format de fichier** et l'**extension** : le format de fichier décrit la façon dont l'information est codée ; l'extension est une partie du nom du fichier.

Quand vous voulez enregistrer un document au format Microsoft Word, le logiciel propose trois formats possibles (Microsoft Word 97/2000/XP, Microsoft Word 95 ou Microsoft Word 6,0). Ce sont trois formats de fichiers différents mais l'extension du fichier sera la même (doc).

L'extension d'un fichier permet au système de lui associer une application.

- Quand on tente d'ouvrir un fichier (en double-cliquant dessus par exemple), c'est l'application associée par défaut à cette extension qui est lancée et ouvre le fichier ;
- L'extension d'un fichier n'est pas obligatoire. Si celle-ci est omise, il faudra certainement indiquer avec quelle application le fichier doit s'ouvrir.

Les extensions de fichiers ne sont pas toujours visibles. Vous pouvez paramétrer votre gestionnaire de fichiers pour qu'elles le soient.

Pour changer le format d'un fichier :

- il ne faut surtout pas le renommer en changeant son extension car cela ne modifiera pas la façon dont l'information a été codée à l'intérieur du fichier ;
- il faut l'ouvrir avec l'application appropriée et l'enregistrer sous un autre nom en changeant le format dans la liste déroulante « type de fichier ».

On distingue deux catégories de fichiers :

- Les **fichiers textes** : les informations sont stockées sous forme de caractères lisibles par tout éditeur de texte.
- Les **fichiers binaires** : les informations sont stockées en binaire et ne peuvent être lues qu'avec le logiciel adéquat.

Attention, les fichiers textes ne concernent pas forcément les documents issus d'un traitement de texte.

Exemples de fichiers textes :

Un classeur au format ODF est constitué d'une archive de fichiers XML (Extensible Markup Language), qui sont des fichiers textes dont l'information est enrichie à l'aide de balises textuelles.

Un fichier au format CSV est un fichier texte qui contient les données d'un tableau séparées par des virgules (séparateur de colonnes) ou des marques de fin de paragraphe (séparateur de lignes). Ce format de fichier peut être exploité comme feuille de calcul dans un tableur.

Une page web au format HTML est un fichier texte utilisant des balises textuelles pour mettre en forme le texte, insérer une image ou définir un lien hypertexte.

Exemples de fichiers de données binaires :

Un document au format Microsoft Word 97/2000/XP.

Une photo au format d'image JPEG.

Un fichier exécutable d'application.

L'interopérabilité

Pour qu'un logiciel puisse exploiter les fichiers produits par un autre logiciel, il doit connaître les spécifications du format des fichiers. Si les spécifications sont publiées par un organisme de normalisation il s'agit d'une **norme** ; si elles sont publiées par un autre type d'organisme, il s'agit d'un **standard**.

Un format est dit **ouvert** si ses spécifications sont publiées et accessibles de tous. Dans le cas contraire, on parle de format **fermé** ou **propriétaire**.

Qu'est ce que l'interopérabilité ?

La **compatibilité** est la possibilité pour deux systèmes de types différents de fonctionner ensemble, notamment en utilisant les mêmes fichiers de données.

Lorsqu'un acteur devient dominant dans un domaine, les autres acteurs font en sorte d'être compatibles avec lui.

- Avantage : l'ensemble des systèmes peuvent à peu près fonctionner ensemble.
- Inconvénient : l'acteur dominant contrôle d'une certaine manière cette possibilité.

L'**interopérabilité** est la possibilité pour différents systèmes de fonctionner ensemble sans dépendre d'un acteur particulier. Elle repose sur la présence d'un standard ouvert.

L'**interopérabilité** est la capacité que possède un produit ou un système, dont les interfaces sont intégralement connues, à fonctionner avec d'autres produits ou systèmes existants ou futurs et ce sans restriction d'accès ou de mise en œuvre.

Explications et définitions adaptées du groupe de travail Interop de l'AFUL (consulté le 16 août 2011).

Exemple :

Le format de fichier Microsoft Word 97/2000/XP est un format fermé. Quand on tente d'enregistrer un document réalisé avec Writer d'Open Office dans ce format, le message suivant apparaît : « Il est possible qu'une partie du formatage et du contenu de ce document ne puisse pas être enregistrée dans le format de fichier Microsoft Word 97/2000/XP. Souhaitez-vous continuer ? »

En effet, OpenOffice ne peut garantir de traiter parfaitement tous les éléments du format Microsoft Word 97/2000/XP car son référentiel n'est pas public.

Les formats de documents

Un **document texte** est un fichier créé par un traitement de texte. C'est un fichier qui contient non seulement des caractères, mais aussi une structure, des mises en forme, des objets, etc.

Voici une liste non exhaustive de formats de fichiers pour les documents texte :

- le format **texte** (txt) est un format de fichiers basique : il est composé d'une suite de caractères simples (l'espace et le retour à la ligne étant considérés comme des caractères) ;
- le format **Microsoft Word** (doc) est un format propriétaire très courant ;
- le format **Rich Text Format** (rtf) est un format ouvert développé par Microsoft dans un souci d'interopérabilité ;
- le format **Texte de l'Open Document Format** (odt) est un format ouvert de traitement de texte ; il est utilisé par l'application Writer d'Open Office ;
- le format **Texte de l'Office Open XML** (docx) est un format de traitement de texte créé par Microsoft pour répondre à une demande d'interopérabilité ; il est utilisé par Microsoft Word depuis 2007.

Un **classeur** est un fichier manipulé par un tableur. Il regroupe plusieurs feuilles de calcul.

Voici une liste non exhaustive de formats de fichiers pour les classeurs :

- le format **Comma-Separated Values** (csv) est un format texte ouvert qui se présente sous forme de lignes de valeurs (sans formules) ; c'est un format d'échange souvent utilisé pour exporter des listes de nombres (liste de notes, etc.) ;
- le format **Microsoft Excel** (xls) est un format propriétaire ;
- le format **Classeur de l'Open Document Format** (ods) est un format de classeur ouvert ; il est utilisé par l'application Calc d'Open Office ;
- le format **Classeur de l'Office Open XML** (xlsx) est un format de classeur créé par Microsoft pour répondre à une demande d'interopérabilité. Il est utilisé par Microsoft Excel depuis 2007.

Pour une présentation en face à face, on accompagne souvent son exposé d'un **diaporama** (suite de diapositives) créé par un logiciel de **présentation**.

Voici une liste non exhaustive de formats de fichiers pour les présentations :

- le format **Microsoft PowerPoint** (ppt) est un format propriétaire ;
- le format **Présentation de l'Open Document Format** (odp) est un format ouvert de présentation ; il est utilisé par l'application Impress d'Open Office ;
- le format **Présentation de l'Office Open XML** (xlsx) est un format de présentation créé par Microsoft pour répondre à une demande d'interopérabilité ; il est utilisé par Microsoft PowerPoint depuis 2007.

Pour une présentation en ligne, on peut créer un site constitué d'une ou plusieurs **pages web**.

Voici un exemple de format de fichier pour les pages web :

- le format de document **HyperText Markup Language** (html) est un format ouvert pour la conception de pages web.

Les formats d'images

Pour représenter une image, on peut la décrire à l'aide de fonctions mathématiques (représentation vectorielle) ou par l'ensemble des points qui la composent (représentation matricielle).

Une **image vectorielle** est définie par un ensemble de données mathématiques : coordonnées, fonctions, attributs, etc.

Une image vectorielle peut être agrandie ou réduite sans dégradation car l'image sera recalculée précisément en fonction de la taille souhaitée. En général, le fichier correspondant est peu volumineux.

Quelques formats d'images vectorielles

- Le format **Scalable Vector Graphics** (svg) est un format ouvert d'image vectorielle ; il est surtout utilisé en cartographie et sur les téléphones portables.
- Le format **Dessin de l'Open Document Format** (odg) est un format ouvert de dessin vectoriel ; il est utilisé par l'application Draw d'Open Office.

Un dessin réalisé avec la barre d'outils dessin des logiciels de bureautique ou un diagramme créé dans un logiciel de présentation sont des images vectorielles.

Une **image matricielle** est définie par une grille de points ou pixels auxquels est associée une couleur.

Une image matricielle se dégrade si on l'agrandit : la pixellisation devient visible. En fonction de la taille de l'image et du nombre de couleurs utilisées, le fichier correspondant peut devenir volumineux. Pour transiter sur Internet, on utilisera des formats matriciels compressés.

Quelques formats d'images matricielles

- Le format **BitMap** (bmp) est un format d'image matricielle développé par Microsoft et IBM. C'est un format simple, ouvert et facilement utilisable par les logiciels mais il est peu utilisé sur le web à cause de la taille volumineuse de ses fichiers, la couleur de chaque pixel étant codée sans effort de compression.
- Le format **Joint Photographic Experts Group** (jpeg ou jpg) est un format d'image ouvert compressé pour des images de type photographie (16 millions de couleurs). Le principe de compression JPEG est de coder la couleur de rectangles dont les pixels ont la même couleur (compression sans perte) ou des couleurs très proches (compression avec perte, taux de qualité paramétrable).
- Le format **Graphics Interchange Format** (gif) est un format d'image ouvert compressé pour des images de type dessin. Il permet de définir une couleur de fond transparente, de proposer un mode d'affichage entrelacé (aperçu de l'image qui se précise en cours de téléchargement) et de créer des images animées. Les couleurs d'une image sont définies sur une palette de 256 couleurs choisies parmi 16 millions de couleurs possibles : c'est le principe de base de la compression GIF, dite en couleurs indexées.
- Le format **Portable Network Graphics** (png) est un format d'image ouvert compressé. Il effectue une compression en couleurs indexées, mais la palette n'est pas limitée à 256 couleurs.

En pratique, on règle le taux de compression de façon à conserver une bonne qualité visuelle de l'image.

En général, la compression JPEG est préconisée pour les photographies, tandis que la compression GIF ou PNG se prête mieux aux dessins.

Les autres formats

Le format **Portable Document Format** (pdf) est un format ouvert développé par la société Adobe Systems qui préserve la mise en forme d'un document :

- il est généralement utilisé pour la diffusion de documents formatés qui n'ont pas vocation à être modifiés ;
- il s'applique à tous types de documents imprimables : document de traitement de texte, feuille de calcul, présentation, etc.

Pour lire un fichier au format PDF, vous devez utiliser un lecteur spécifique : Adobe Reader est un lecteur gratuit diffusé par la société Adobe Systems.

La numérisation du son et de la vidéo permet de stocker des musiques et des films sous forme de **fichiers**.

Étant donné la taille importante des fichiers audio et vidéo, il est fortement conseillé d'utiliser des formats audio-visuels compressés.

- Le format **Windows Media Audio** (wma) est un format propriétaire de compression audio développé par Microsoft.

Le **Moving Pictures Expert Group** a développé des méthodes de compression pour les fichiers audio et vidéo.

Parmi les plus connus, on distingue :

- le format MPEG-1/2 Audio Layer 3 ou **MP3** (mp3) est un format compressé ouvert pour les fichiers audio ; il s'agit d'une compression avec des pertes non perceptibles par l'oreille humaine ;
- le format **MPEG-4** (mpg, mp4, etc.) est un format compressé ouvert pour les fichiers vidéos.

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 mébiocet (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 tébiocet (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.