

Format de Fichier

Une introduction
Simon CHOLLET

- Introduction
- Les différents formats
- TP





Introduction

- Format de fichier : qu'est-ce que c'est ?
- Principaux formats
- Nommer les fichiers



Format de fichier : qu'est-ce que c'est ?

- Série de **données** stockées dans un **fichier**
- Chaque logiciel a sa **propre manière** de coder les informations
- Les fichiers peuvent être de **2 types** :
 - **ASCII** : on peut l'ouvrir avec un éditeur de texte
 - **Binaire** : illisible ... à part passer par des éditeurs spécialisés
- Il existe plusieurs **modes** :
 - **Open Source** : la façon dont il est composé est connue
 - **Propriétaire** : il faut le logiciel particulier pour l'ouvrir
- Un **format** de fichier : c'est pour **représenter** des données
- Représente : texte, page, image, son, exécutable, etc.
- Permet d'**échanger** des données entre divers programmes informatiques ou logiciels
- **Interopérabilité** : possibilité d'échanger des données entre différents logiciels





Principaux formats

- Documents **textuels** : catégorie la moins bien normalisée
 - Principalement du texte, quelques illustrations
- **Images** : catégorie très bien normée
 - Photos, illustrations, dessins : format **vectorel** ou **matriciel**
 - Informations supplémentaires « méta données »
- **Audios** : catégorie très bien normée
 - Enregistrements, musiques
 - Informations supplémentaires « méta données »
- **Vidéos** : catégorie très bien normée
 - Films, montages
 - Informations supplémentaires « méta données »
- **Compression – Archivage** : catégorie très bien normée
 - Réduction de l'espace occupé
- **Exécutables**, etc.





Nommer les fichiers

- Important de « bien » **nommer** les fichiers
- **Règles** : éviter caractères spéciaux, éviter PB de transmission
- **Extension** (3 lettres) de fichier = **son type**
 - nom_de_fichier.ext
 - Exemple : cv-2020.rtf
- **Quelques** extensions :

Extension	Type de fichier
.doc	Document Word
.rtf	Texte mis en forme (Rich Text Form)
.jpg	Image au format JPEG
.mov	Vidéo au format Quicktime
.mp3	Vidéo (ou son) au format MPEG 3
.png	Image au format Portable Network Graphics
.gif	Image au format Graphics Interchange Format
.html	Page web





Quelques formats

- Le texte
- Les images
- Les vidéos
- Les sons
- Compression, archivage
- Les autres ...



Le texte (1/2)

- **DOC, DOCX** (binaire) :
 - Fichier document de Microsoft Word
- **HTML, HTM, SHTM, SHTML** (ASCII) :
 - **H**yper**T**ext **M**arkup **L**anguage
 - Page Web
- **PDF** (ASCII ou binaire) :
 - **P**ortable **D**ocument **F**ormat
 - Peut être **crypté, signé**
- **PPT, PPS** (binaire) :
 - **P**ower**P**oint Presen**T**ation, **P**ower**P**oint **S**lide Show
- **RTF** (ASCII) :
 - **R**ich **T**ext **F**ormat
 - Mise en page **possible**





Le texte (2/2)

- **ODT** (binaire) :
 - **O**pen **D**ocument **T**ext
 - Spécifique à Open Office ou Libre Office
- **ODP** (binaire) :
 - **O**pen **D**ocument **P**resentation
 - Spécifique à Open Office ou Libre Office
- **TXT** (ASCII) :
 - Texte « pur » : **pas** de mise en page possible
- **SGML**, **XML** (ASCII) :
 - **S**tandard **G**eneralized **M**arkup **L**anguage, **eX**tensible **M**arkup Language
 - Fichier texte **hiérarchisé**





Les images (1/3)

- **BMP** (matriciel) :
 - **BitMaP**
 - Image non compressée, natif de Windows
- **DXF** (vectoriel) :
 - Dessin de **CAO** de AutoDesk pour Autocad
- **PSD** (matriciel) :
 - Natif du logiciel **Photoshop**
 - Image non compressée
- **AI** (vectoriel) :
 - Natif du logiciel **Adobe Illustrator**
- **EPS** (matriciel ou vectoriel) :
 - **Encapsulated PostScript**
 - Contenu image / page complète : textes, filets, images importées
 - Ne peut contenir qu'une page





Les images (2/3)

- **PS** (vectoriel) :
 - **P**ost**S**cript
 - Format transmissible directement aux serveurs d'impression
- **GIF** (matriciel) :
 - **G**raphics **I**nterchange **F**ormat
 - Utilisé pour produire de **petites images** (256 couleurs)
 - **Transparence** et **animations** disponibles
- **JPEG** (matriciel) :
 - **J**oint **P**hotographic **E**xperts **G**roup
 - Image **compressée** sans grosse perte de données
 - Pour les photos : **16.10⁶** couleurs
- **PICT** (vectoriel) :
 - Interne au Macintosh





Les images (3/3)

- **TIFF, TIF** (matriciel) :
 - **T**agged **I**mage **F**ile **F**ormat
 - Pour l'**archivage** d'images de qualité
 - Algorithme de compression **sans perte**
- **PNG** (matriciel) :
 - **P**ortable **N**etwork **G**raphics
 - **Norme publique**, voué à remplacer le GIF (non libre)
 - **Transparence** disponible
- **PCX** (matriciel) :
 - Issu du logiciel Paintbrush sous Windows
- **SVG** (vectoriel) :
 - **S**calable **V**ector **G**raphics





Les vidéos

- **AVI** :
 - **A**udio **V**ideo **I**nterleave
 - Format vidéo générique pour **Windows**
 - Compression **possible** par codec (codeur-décodeur)
- **MOV, QT** :
 - QuickTime Movie
- **MPEG, MPG** :
 - **M**oving **P**ictures **E**xpert **G**roup
 - Norme **publique** : format ouvert / Open Source
 - Très répandu





Les sons

- **AIFF, AIF** :
 - **A**udio **I**nterchange **F**ile **F**ormat
- **AU** :
 - Développé pour UNIX
- **MIDI, MID** :
 - **M**usic **I**nstrument **D**igital **I**nterface
 - Norme pour **transmission** informations entre instruments
- **MP3** :
 - **M**oving **P**ictures **E**xpert **G**roup
- **OGG** :
 - Format libre
- **WAV** :
 - **W**aveform Audio File Format





Compression, archivage

- **GZ, GZIP** :
 - GNU ZIP
 - Sur la plateforme **UNIX**
 - Utilisable aussi sous Windows
- **RAR** :
 - Roshal **AR**chive
 - Efficace sur de très gros fichiers
 - Permet le **morcellement** en plusieurs fichiers
- **TAR** :
 - Tape **AR**chive
 - **Pas** de compression, utilisé souvent avec **GZ**
- **ZIP** :
 - Format de fichier **commun**





Les exécutables

- **EXE** :
 - Spécifique à **Windows**
- **BAT** :
 - Spécifique à Windows, **MS-DOS**
 - Script en **mode texte** avec suite d'opérations à exécuter
- **APPIMAGE** :
 - Spécifique Linux
- **SH** :
 - Spécifique à **Linux**
 - Script en **mode texte**, avec suite d'opérations à exécuter
- **ELF** :
 - Executable and **L**inking **F**ormat
- **DLL, LIB** :
 - **Librairies** de fonctions





Les autres ...

- **RA, RAM** : **R**eal **A**udio :
 - Diffusion, **streaming** sur Internet
 - Propriétaire
- **SWF** : Shockwave Flash
- **XLS, XLSX** : Excel, spécifique Windows
- **INI, CONF, CNF** : Fichier texte de **configuration**
- **JS** : **J**ava**S**cript
- **JSON** : **J**ava**S**cript **O**bject **N**otation
- **CSS, PHP** : Sur le Web.
- Etc ...





TP : on pratique !

- Propriétés de fichiers
- Transformation PDF → JPG
- Compression de fichiers / répertoire
- Décompression
- Découpage PDF

Tests



- Changer l'extension d'un fichier '.doc' en '.jpg' et essayer de l'ouvrir.
- Regarder les propriétés des images, des différents fichiers.
- Transformer un fichier PDF en fichier JPEG, ou PNG, avec GIMP.
- Compresser des fichiers JPEG, des fichiers BMP, un répertoire.
- Tester la décompression.
- Tester le traitement de fichier PDF avec « pdfarranger »



FIN !

Merci ...

- ... pour avoir tenu jusqu'ici,
- ... de ne pas hésiter à poser des questions !

