# i100 - Distinguer, préparer et évaluer des données

Une image contenant objet

Description générée automatiquementLedda Anthony

[i100 - Distinguer, préparer et évaluer des données 1](#_Toc23758816)

[1. Définition du titre 3](#_Toc23758817)

[2. Qu’est-ce qu’un format 3](#_Toc23758818)

[3. Qu’est-ce qu’un format de données 3](#_Toc23758819)

[4. Code ASCII 4](#_Toc23758820)

[5. Format des numéros de téléphone 5](#_Toc23758821)

[6. Format des adresses e-mail 6](#_Toc23758822)

[7. Qu’est-ce qu’un fichier 6](#_Toc23758823)

[8. Format d’une URL 7](#_Toc23758824)

[9. Charset 8](#_Toc23758825)

[10. QWERTZ 8](#_Toc23758826)

[11. Unicode 8](#_Toc23758827)

[12. Changer la langue d’affichage Windows (fichier .cab): 9](#_Toc23758828)

[13. Alt 124 10](#_Toc23758829)

[14. Git et Github 11](#_Toc23758830)

[15. Données biométrique : Passeport suisse et Carte d’identité 12](#_Toc23758831)

[16. Signature de fichier 13](#_Toc23758832)

[17. Scan code 14](#_Toc23758833)

[18. Métadonnées 15](#_Toc23758834)

# Définition du titre

Distinguer : Par exemple différencier le nom de famille et le prénom d’un utilisateur. C’est reconnaître, faire la différence entre plusieurs éléments.

Préparer : Dans notre thème ce sera « clavioter » des données dans un formulaire par exemple. C’est organiser et structurer un ou plusieurs éléments, via un tableur ou une base de données.

Évaluer : Vérifier par exemple que le mot « Jean » est bien le prénom et non le nom. C’est analyser, tester ou juger quelque chose.

Données : Une date de naissance, un nom et un prénom sont des données propres à chaque individu. En informatique, les données représentent des informations dans un programme. Elles peuvent se trouver sous différentes formes : textes, numérique, images et sons.

Source : https://fr.wikipedia.org/wiki/Donn%C3%A9e\_(informatique); <https://tools.ietf.org/html/rfc8610> ; connaissances personnelles

# Qu’est-ce qu’un format

C’est la façon dont laquelle est structurée quelque chose.

Source : https://www.larousse.fr/dictionnaires/francais/format/34637

# Qu’est-ce qu’un format de données

C’est la façon dont est représenté un type de données (valeur que peut prendre une donnée), codé en binaire. Il est dit que la donnée est représentée comme un nombre.

Il existe plusieurs types de formats de données :

* Formats de nombres
* Formats de texte
* Formats d’image
* Formats de vidéo
* Formats de scène 3D
* Formats de son
* Formats classiques

Les caractères de la table ascii (ci-dessous) est codé en bits.

Par exemple « E » sera codé 100 0101, soit 69 en décimal.

Source : <https://fr.wikipedia.org/wiki/Format_de_donn%C3%A9es>

# Code ASCII

(USASCII) ASCII : USA Standard Code for Information Interchange (7bits)

Ce code est utilisé sur des connections hôte à hôte (clients ou serveurs qui envoient ou reçoivent des données).

Sur le tableau la représentation bit de « A », par exemple, sera :

**Pour composer un caractère ASCII avec le clavier** il faut appuyer simultanément sur « alt » et le nombre qui représente le symbole dans la table ASCII en base de dix.

Une image contenant texte

Description générée automatiquementExemple pour « A » : **65** en base de dix **« alt + 65 »** **A**

Source : <https://tools.ietf.org/html/rfc20>

# Format des numéros de téléphone

00/011/+ 41 (0)21 765 43 21

Code d’accès international Indicatif international Indicatif régional Numéro de téléphone local

Un numéro de téléphone est une suite de chiffres ou de lettres, leur structure varie selon les pays.

Chaque numéro possède deux formats :

* Un pour les appels internes au pays : 021 765 43 21
* Un pour les appels internationaux : +41 21 765 43 21 Exemple numéro téléphone fixe Lausanne

## Un numéro de tél. contient :

* Un code d’accès internationalqui doit être composé pour les appels à l’étranger. Depuis un téléphone fixe, le préfixe « 00 » dans les pays européens, « 011 » dans les pays américains et un « + » depuis un mobile.
* **Un** [**indicatif international**](https://fr.wikipedia.org/wiki/Liste_des_indicatifs_t%C3%A9l%C3%A9phoniques_internationaux_des_r%C3%A9seaux_mobiles), un préfixe qui indique le pays, il doit être ajouté avant le numéro pour joindre un appel à l’étranger. Par exemple : « 41 » pour la suisse.
* **Un** [**indicatif régional**](https://wiki.epfl.ch/voip-at-home/indicatifs-regions-suisse) **pour les téléphone fixe**,un préfixe qui indique la région, ce sont les 3 (2 depuis l’international) premiers chiffres du numéro local fixe. Par exemple : « 021 » pour Lausanne.
* Un indicatif mobile pour les mobiles, un préfixe distribué par l’opérateur mobile. Par exemple : « 075 » pour Swisscom
* Un numéro de téléphone local, c’est le numéro qui sera attribué à un particulier ou une entreprise. En suisse ces numéros commencent par un « 0 », qu’il faudra composer seulement si l’appel est local.

## Syntaxe numéro de tél. :

[États-Unis](https://fr.wikipedia.org/wiki/%C3%89tats-Unis) : +1 (CCC) CCC-CCCC 10 chiffres au format national, 11 chiffres au format international.



[Suisse](https://fr.wikipedia.org/wiki/Suisse) : +41 CC CCC CC CC 10 chiffres au format national, 11 chiffres au format international.



[Drapeau du Brésil](https://commons.wikimedia.org/wiki/File:Flag_of_Brazil.svg?uselang=fr)[Brésil](https://fr.wikipedia.org/wiki/Br%C3%A9sil) : +55 CC CCCC-CCCC 10 chiffres au format national, 12 chiffres au format international

Source : <https://fr.wikipedia.org/wiki/Num%C3%A9ro_de_t%C3%A9l%C3%A9phone> <https://www.indicatifs.fr/suisse>

# Format des adresses e-mail

Jean.pierre @ gmail.com

Partie locale Caractère séparateur Adresse du serveur (domaine)

Une adresse électronique est une suite de caractères qui permet la conduite de courriers électronique dans des boîtes aux lettres informatiques.

## Une **adresse e-mail** contient :

* Une partie locale, c’est l’identifiant, en général un nom d’une personne ou d’un service. Par exemple : Jean.pierre
* Le caractère séparateur @,signifiant at (« à » ou « chez ») en anglais
* L’adresse du serveur, c’est généralement un nom de domaine qui identifie l’entreprise où se trouve la boîte électronique. Par exemple : « gmail.com » pour google mail

## Syntaxe d’une adresse :

* Une adresse doit contenir un « @ (valeur ASCII 64)».
* Les caractères avant et après l’arobase ne doivent pas être un point (« . »).
* Le point ne doit pas apparaître deux fois de suite.
* Elle formée de caractères US-ASCII (p.3)
* L’ID (partie locale) ne peut pas contenir plus de 64 caractères.

Source : <https://fr.wikipedia.org/wiki/Adresse_%C3%A9lectronique> <http://abcdrfc.free.fr/rfc-vf/pdf/rfc5322.pdf>

# Qu’est-ce qu’un fichier

C’est un ensemble de données numériques groupées sous un même nom

Chaque fichier comporte :

* Un nom, il sert à désigner le contenu
* Des métadonnées, tel que : une longueur de fichier, un auteur, une date
* Une extension (liste des principales extensions ci-dessous)

Une image contenant capture d’écran

Description générée automatiquement

Source : https://fr.wikipedia.org/wiki/Fichier\_informatique

# Format d’une URL

http(s):// user :password@ [www.ccm.net](http://www.ccm.net) :80 /glossair/glossair.php

Protocole identifiant et mot de passe (facultatif) Adresse du serveur (domaine) Port (facultatif si 80) Chemin

« Une **URL** (*Uniform Resource Locator*) est un format de nommage universel pour désigner une ressource sur Internet. »

Citation : <https://www.commentcamarche.net/contents/542-qu-est-ce-qu-une-url>

## Une adresse url contient :

* Un protocole, Il y a de nombreux protocoles utilisés (HTTP, FTP, News, Mailto, Gopher, …). Le plus utilisé est le protocole HTTP (HyperText Transfer Protocol), il permet d’échanger des pages Web au format HTML.

HTTPS permet de vérifier l'identité du site web.

* **Un identifiant et un mot de passe (facultatif),** ils permettent de sécuriser l’accès au serveur, cette option est déconseillée car le mot de passe est visible dans l’URL.
* **Un nom du serveur,** c’est le domaine de la machine qui héberge la ressource demandée, le nom de serveur peut être remplacé par l’adresse IP.
* **Un port (facultatif),** c’est un numéro qui permet au serveur de savoir quelle type ressource est demandée. **Le port par défaut du protocole est le port 80**.
* **Un chemin,** C’est ce qui permet au serveur de connaître l’emplacement où la ressource est située.

## Syntaxe d’une URL :

* Techniquement, il n’y a pas de limite à la taille de l’URL, mais en pratique chaque navigateur et serveur web ont leur propre limite.
* Dans internet explorer la limite de caractères d’une URL est de 2083 caractères.
* La taille maximale d’un nom de domaine est de 255 caractères.
* Il ne peut y avoir de caractères accentués et spéciaux.

Source : <https://www.commentcamarche.net/contents/542-qu-est-ce-qu-une-url> <https://www.webrankinfo.com/dossiers/conseils/longueur-url-seo>

# Charset

**Définition :** Charset vient de « character set », qui veut dire « ensemble de caractères ». L’ASCII et l’Unicode sont deux charsets très connus.

Source : <https://www.yakaferci.com/definition/charset/24>

# QWERTZ

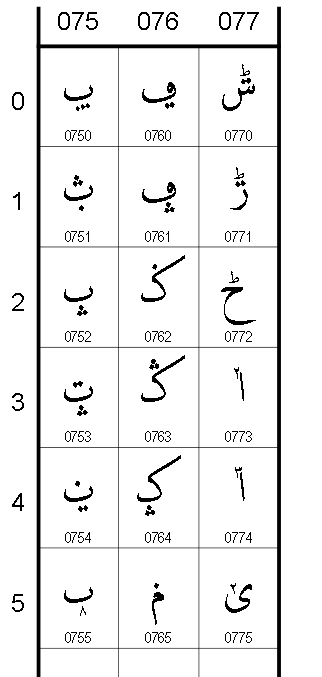
Le clavier QWERTZ est une disposition de touches germanophone (Allemagne, Autriche, Suisse).

Source : https://fr.wikipedia.org/wiki/QWERTZUne image contenant équipement électronique

Description générée automatiquement

Figure - https://fr.wikipedia.org/wiki/QWERTZ#/media/Fichier:KB\_Swiss\_Luxemburg\_text.svg

# Unicode

Unicode permet l’échange de textes dans différentes langues, à l’échelle mondiale. Par défaut codé en UTF-8.(8 bits)

**Comment faire des signes étranges trouvés dans UNICODE avec un clavier QWERTZ  :**

* Aller sur <http://www.unicode.org/charts/>
* Choisir un character code chart
* Se référer au tableau du PDF (ex : [Arabic Supplement](http://www.unicode.org/charts/PDF/U0750.pdf) à droite)
* Pour le glyphe « ݐ » le code sera 0750 en hexadécimal
* Taper (le code en hexadécimal) +alt + C
* 0750 + alt + C => « ݐ »

Source : https://fr.wikipedia.org/wiki/Unicode

Figure - https://www.unicode.org/charts/PDF/U0750.pdf

# Changer la langue d’affichage Windows (fichier .cab):

Télécharger une langue d’affichage sur : <https://www.tobias-hartmann.net/2016/11/windows10-mui-language-pack-1607-build-14393-download/>

Appuyer simultanément sur **Win + R** pour ouvrir la fenêtre de dialogue

Clavioter « lpksetup.exe » et presser Enter

Cliquer sur installer des langues d’affichage

Une image contenant capture d’écran

Description générée automatiquementCliquer sur parcourir, sélectionner votre fichier .cab et cliquer sur suivant

Une image contenant capture d’écran

Description générée automatiquementAller dans paramètres>Heure et langue>langue>langue d’affichage

Source : <https://winaero.com/blog/how-to-install-an-mui-language-cab-file-in-windows-10/>

Pour téléchargement languages pack <https://www.tobias-hartmann.net/2018/12/windows-10-language-packs-1809-build-17763-oktober-2018-update-download/>

# Alt 124

Le code alt 124 sous ASCII donne le symbole « | » et correspond au nombre Unicode U+007C.

|  |  |  |
| --- | --- | --- |
| Décimal | Nombre Unicode |  |
| …. |  |  |
| 121 | U+0079 | y |
| 122 | U+007A | z |
| 123 | U+007B | { |
| 124 | U+007C | | |
| 125 | U+007D | } |
| 126 | U+007E | ~ |
| 127 | U+007F | DEL |

Une image contenant capture d’écran

Description générée automatiquement

**Si on ne possède pas de pavé numérique**, il est quand même possible d’écrire des caractères en code ASCII en utilisant un clavier numérique, pour cela il faut presser : **Win + ctrl + O**

Ensuite sur le clavier virtuel allez dans option et cochez activer le pavé numérique.

Pour faire un caractère avec le code alt suivi de code ascii en décimal, tenez enfoncé **alt + fn** PHYSIQUEMENT et cliquez avec la souris sur les **chiffres dans le pavé numérique du clavier virtuel**.

Une image contenant clavier, équipement électronique, intérieur, photo

Description générée automatiquement

**PHYSIQUEMENT**

Source : <https://support.microsoft.com/fr-fr/help/10762/windows-use-on-screen-keyboard>

# Git et Github

On utilise un système de versioning pour sauver nos données en informatique.

Tutoriel git et GitHub :

* **Se connecter sur GitHub**
* **Créer un nouveau dépôt**
* **Installer git sur le pc**
* **Définir un user Name et une adresse email :**

$ git config --global user.name "John Doe"

$ git config --global user.email [johndoe@example.com](mailto:johndoe@example.com)

* **Cibler un fichier/dossier :** $ cd module\_100
* **Initialiser git sur le fichier ciblé :** $ git init
* **Pointer vers le dépôt distant :** $ git remote add i100 https://github.com/anthonyledda/module100.git
* **Clonez le dépôt local :** $ git clone <https://github.com/anthonyledda/module100.git>
* **Upload le document Word sur GitHub**
* **Valider le commit**
* **Pour importer fichier :** $ git pull <https://github.com/anthonyledda/module100.git>

Une image contenant capture d’écran

Description générée automatiquementSource : <https://openclassrooms.com/fr/courses/5641721-utilisez-git-et-github-pour-vos-projets-de-developpement/6113016-installez-git-sur-votre-ordinateur>

# Données biométrique : Passeport suisse et Carte d’identité

## Une image contenant texte Description générée automatiquementUn document d’identité suisse contient :

* Nom d'état civil;
* Prénoms;
* Sexe;
* Date de naissance;
* Lieu d'origine;
* Nationalité;
* Taille;
* Signature;
* Photographie;
* Autorité d'établissement;
* Date d'établissement;
* Date d'expiration;

Figure 3 http://www.biwi.ch/assets/thumbs/135/1457419986\_carte\_identite.9d6ca4d7.jpg

* Numéro et type du document.

## Avec une demande spéciale le document peut aussi contenir :

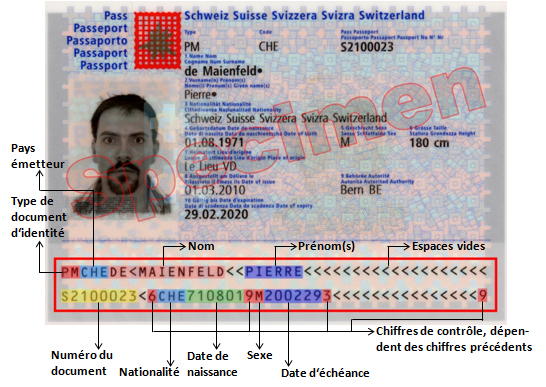
* Le nom d'alliance
* Le nom reçu dans un ordre religieux
* Le nom d'artiste
* Le nom de partenariat
* Pour les mineurs : le nom de leurs représentants légaux.

## … et la mention de signes particuliers :

* Handicaps
* Prothèses
* Implants

## Machine Reading Zone

Sur les documents d’identité il existe un code conçu pour les machines et non lisible par des humains.

****Sur cet exemple (un passeport suisse) :

* **Le type de document :** 2 lettres (PM, ID, TS, *Catégories permis de conduire* : A1, …)
* **Pays émetteur :** 3 lettres (CHE, FRA, …)
* **Nom :** 45 caractères
* **Prénom :** 45 caractères =>passeport, 30 caractères =>carte d’identité
* **\*Espaces vides**
* **Numéro du document :** 1 lettre + 7 chiffres
* **Nationalité :** 3 lettres (CHE, FRA, …)
* **Date de naissance :** 6 chiffes (AA.MM.JJ)
* **Sexe :** M ou F
* **Date d’échéance :** 6 chiffres (AA.MM.JJ)
* **Chiffres de contrôle :** 4 chiffres réparti dans le code

Figure 4 https://blocnotes.iergo.fr/articles/recherche-utilisateur-fait-maison/

Source : <https://www.admin.ch/opc/fr/official-compilation/2010/607.pdf>

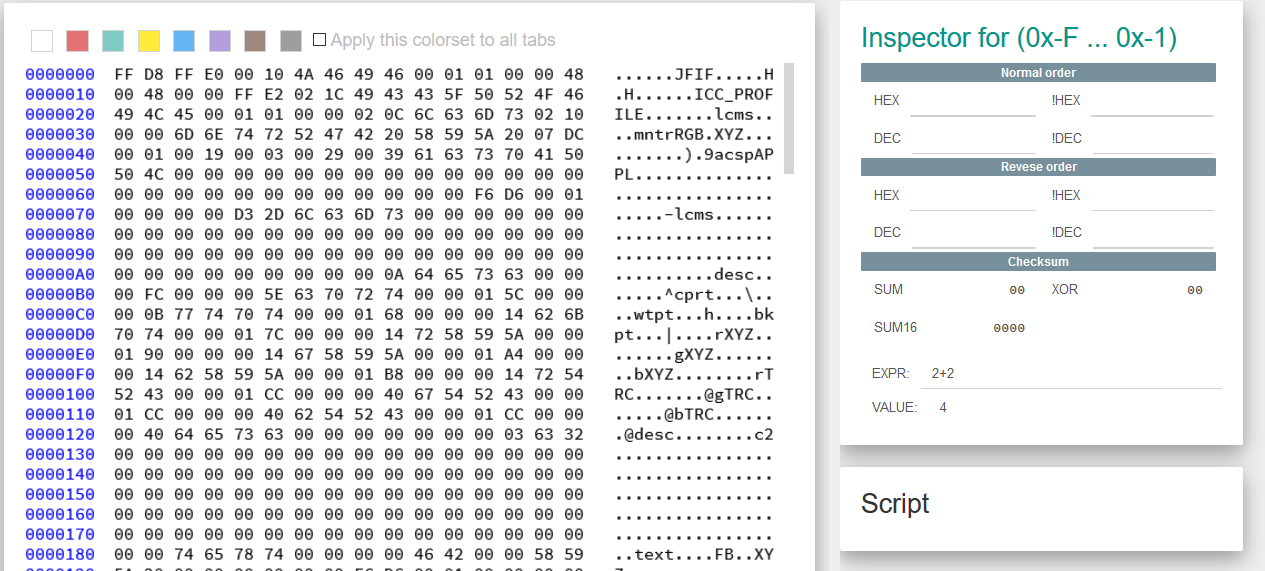
<https://www.admin.ch/opc/fr/classified-compilation/19994375/index.html>

# Signature de fichier

On analyse un fichier avec un éditeur hexadécimal (ici : <https://hex-works.com/eng>)

Chercher une photo aléatoire sur Google (par exemple : 22.09.2019)

Figure 5 http://www.alsacemonde.org/wp-content/uploads/F%C3%AAte-des-vendanges-Montr%C3%A9al.jpg

* Aller sur un éditeur hexadécimal en ligne (<https://hex-works.com/eng>)
* Dans l’éditeur ouvrir le fichier
* Reporter les premiers caractères (pour cet exemple 24 caractères)
* Se référer à la liste des signatures de fichiers Wikipédia : <https://en.wikipedia.org/wiki/List_of_file_signatures>
* Ici « FF D8 FF E0 00 10 4A 46 49 46 00 01» correspond à un fichier \*.JPG ou \*.jpeg

**Quelques exemples de signatures de fichiers :**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 52 49 46 46 ?? ?? ?? ?? 41 56 49 20 | |  | | --- | | RIFF....  AVI. | | avi | Audio Video Interleave video format |
| FF FB | ÿû | mp3 | MPEG- 1 Layer 3 file without an ID3 tag |
| 25 50 44 46 2d | %PDF- | pdf | PDF document |

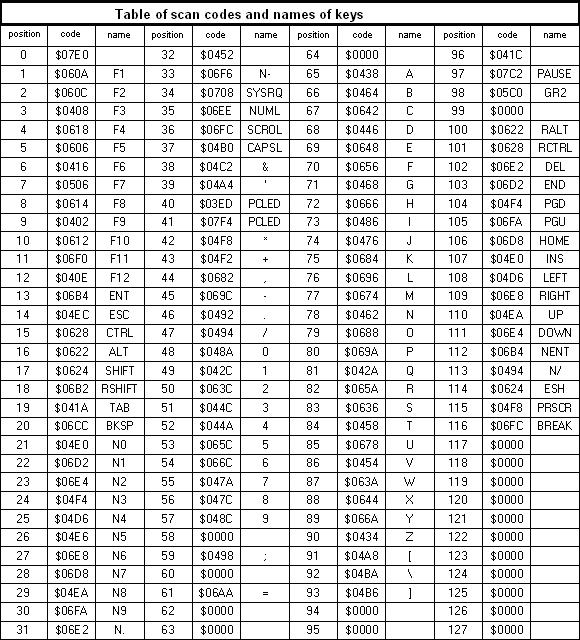
Source : <https://en.wikipedia.org/wiki/List_of_file_signatures>

# Scan code

Un scan code est la donnée que le clavier envoie au PC pour transmettre quelle touche a été pressée.

Un chiffre ou un nombre correspond à chaque touche du clavier.

Lors de la pression de la touche, l’un des commutateurs du clavier s’active et le clavier stocke sa valeur dans une petite mémoire cache en informant l’ordinateur. Ensuite l’ordinateur lit le cache et prend les mesures nécessaires.

Figure 6 https://external-content.duckduckgo.com/iu/?u=https%3A%2F%2Fhw-server.com%2Fimages%2Fcodes.jpg&f=1&nofb=1

Source : <https://www.cknow.com/cms/articles/what-is-a-scan-code.html>

<https://en.wikipedia.org/wiki/Scancode>

# Métadonnées

Une métadonnée défini une autre donnée. Il précise des informations sur un certain élément

Par exemple une image possède des métadonnées sur sa largeur, sa date de création, parfois la localisation lors de la prise de la photo.

## Exif

Exif ou Exchangeable image file format est une spécification de format qui vient ajouter des balises métadonnées.

Les balises métadonnées définies par l’EXIF comporte :

* L’information de la date et de l’heure
* La marque et le modèle de l’appareil
* Des informations géographiques
* Les informations sur les droits d’auteur

Ligne 1240 correspond à la date de prise de vue

Source : <https://fr.wikipedia.org/wiki/Exchangeable_image_file_format>

<https://techterms.com/definition/metadata>