

Les formats de fichier numérique : les grands principes

1. Qu'est-ce qu'un fichier numérique ?

Un fichier est un ensemble de données numériques qui :

- porte une **extension en 3 lettres** en y précisant le **format**,
- est enregistré sur un **support de stockage**,
- est manipulé comme une **unité**.

2. Qu'est-ce qu'un format de fichier ?

Le format de fichier est **une des métadonnées techniques** constitutive du fichier numérique. Il désigne la **nature d'un document informatique** et permet **d'identifier le logiciel** nécessaire à sa lecture.

3. Comment identifier un format de fichier ?

Un format de fichier s'identifie en fonction des trois critères suivants :

- **Extension** : suite de caractères à la suite d'un nom de fichier (exemple : .odt, .xls, .pdf, ...),
- **Type MIME** : standard permettant d'indiquer la nature et le format d'un document.
- **Signature** : ensemble de caractères propres à un format de fichier et permet de spécifier le format ou le protocole utilisé.

4. Comment garantir la pérennité d'un format de fichier ?

Cinq critères permettent de **définir le niveau de pérennité** d'un format :

- **Spécificité** : il existe une description de la convention utilisée pour représenter la donnée et celle-ci est suffisamment décrite pour en développer une implémentation complète.
- **Ouvert** : la convention est publique (sinon le format est dit fermé). Elle est donc sans restriction d'accès ni de mise en œuvre.
- **Normalisation** : la convention est adoptée par des organismes de normalisation tels que ISO ou W3C.
- **Standardisation** : il n'existe pas de norme sur ce format mais son utilisation est tellement répandue qu'il est considéré comme un standard.

- **Propriétaire** : si l'exploitation du format entre dans le cadre du droit privé, il dépend alors de l'existence du propriétaire. Il peut alors être publié : le .pdf ou .odt.