



Des formats de fichiers universels

publié le 01/09/2015 - mis à jour le 15/06/2017

Descriptif :

Le choix de formats ouverts facilite les échanges de documents.

La diversité de formats de fichiers répond parfois à des choix techniques justifiés, mais reflète plus souvent une guerre économique souterraine que se livrent les éditeurs de logiciels.

Il existe des formats ouverts ou semi ouverts et interopérables que nous devons privilégier face aux formats dits propriétaires et qui sont restrictifs.

Ces formats de texte interopérables sont :

- **ODT**, Open Document Texte d'Open Office - Libre Office - Neo Office (les suites bureautiques libres en général), mais parfaitement géré par Microsoft Word tant en enregistrement qu'en lecture. ODT peut même être le format d'enregistrement par défaut de Microsoft Word. Open Document Texte est le format de fichier le plus ouvert et interopérable possible à l'heure actuelle.
- **TXT**, c'est le texte brut, aucune marque de mise en forme ne s'y trouve. Ni police, ni gras... Les seuls caractères, espaces, ponctuation et paragraphes. L'avantage de ce format est sa légèreté numérique et sa rapidité de manipulation. Il pourra être très intéressant pour assembler des textes ou en importer dans d'autres applications de production de contenu.
- **PDF**, Ce format est un format ouvert qui préserve la mise en forme d'un fichier telle qu'elle a été définie par son auteur, et cela quels que soient le logiciel, le système d'exploitation et l'ordinateur utilisés pour l'imprimer ou le visualiser.

http://fr.wikipedia.org/wiki/Portable_Document_Format

Le Portable Document Format (PDF) peut aussi être utilisé comme un simple conteneur d'images et de ce fait devenir un piège. Le cas le plus courant est la numérisation d'une page de manuel avec un scanner ou un photocopieur qui va produire un fichier PDF. Dans ce cas le texte n'est pas numérique. Il faut alors traiter ce fichier image par reconnaissance optique pour en extraire le texte.

- Plus d'informations sur les Standards Ouverts et ce qui les définit : documentfreedom.org