

Quels critères applicables aux formats pour la pérennisation ?

Présentation Groupe PIN
15 janvier 2003

Claude Huc

Pérennisation : rappel sur les approches proposées

➤ Migration logiques

- à chaque nouvelle version de Word, relire les fichiers et les ré-enregistrer avec la nouvelle version et contrôler la nouvelle version, tâche non automatique d'une ampleur croissante
- totale dépendance par rapport au propriétaire du logiciel

➤ Emulation

➤ Ordinateur virtuel universel

- créer des informations accompagnées de leur information de représentation de telle sorte telle sorte qu'on ne dépende plus ni d'un matériel ni d'un logiciel particulier pour les lire, maintenant ou plus tard
- Approche OAIS**

Terminologie

➤ **format standard** : défini par un organisme national (AFNOR) ou international (ISO) de normalisation

- le consortium W3C (World Wide Web Consortium) peut aussi être considéré comme tel
- organismes ou comités de normalisation propres à des domaines professionnels (ex : terminologie archivistique)
- attention : un standard peut n'être qu'un cadre à l'intérieur duquel doivent être insérées des éléments qui peuvent ou non être standards, voire privés

➤ **format propriétaire**

- défini par une entreprise ou un propriétaire privé
- protégé par des droits (propriété intellectuelle, copyright)
- publié ou non publié

➤ **formats généralistes ou format métier**

Groupe PIN - 15 janvier 2003

Claude HUC - CNES

3

Critères par rapport à la conservation Le cas idéal ?

- Indépendant de tout propriétaire
- Reconnu au niveau international (standardisation)
- Convertible à tout moment vers des formats de présentation d'usage courant (forme papier, publication sur le Web)
- Lisible à l'oeil nu à l'aide d'un simple éditeur (cas du texte)
- Facile à générer (mesurables ?)

Groupe PIN - 15 janvier 2003

Claude HUC - CNES

4

- Un format de données pourra suivant les cas être :
 - recommandé - format standard approprié par rapport à l'objectif visé
 - acceptable
 - ❖ format standard non directement approprié,
 - ❖ format propriétaire publié - PDF
 - déconseillé :
 - ❖ format propriétaire avec absence d'autre solution, donc solution d'attente
 - inacceptable : format non publié
 - ❖ Word...
- Portabilité/neutralité

un format sera qualifié de neutre s'il n'est pas spécifique d'un système d'exploitation, d'un logiciel ou d'un matériel particulier
- Ces critères sont différents des critères par rapport à l'échange de données (faibles coût des supports et équipement, matériels largement répandus)

- Format propriétaire non publié : a priori inacceptable
 - pire encore lorsqu'il y a incompatibilité d'une version à l'autre
- Format propriétaire publié : peut être une solution d'attente qui reste dangereuse à long terme
 - la politique commerciale du propriétaire peut changer
- Format lisible sur toutes les machines ? (portabilité) Si ce n'est pas le cas, c'est le signe d'une dépendance par rapport à un environnement
- Mais aussi :
 - quelle pourra être la richesse et la pertinence des moyens de recherche de l'information archivée ?
 - ❖ Sur le titre et l'auteur seulement ?
 - ❖ Sur la structure du document ?
 - ❖ Sur telle ou telle partie de son contenu ?
 - interaction forte avec l'utilisation plus ou moins grande des archives



Critères retenus dans la norme NF Z 42-013

- Compatibilité ascendante des différentes versions
- spécifications publiées et accessibles librement
(mais on ne dit pas utilisables librement ?)
- possibilité de représenter des documents mono ou multipage dont le contenu peut être une combinaison de texte, d'image et de graphique
mérite explication