

Question 5	La numérisation suffit-elle à valoriser l'information ?	
Chapitre 5.4	<i>Les métadonnées</i>	Bac STMG SIG
Cours		Lycée J. Feyder

Chapitre 5.4

Les métadonnées

Question 5	La numérisation suffit-elle à valoriser l'information ?	
Chapitre 5.4	<i>Les métadonnées</i>	Bac STMG SIG
Cours		Lycée J. Feyder

1° Définition d'une métadonnée

Une métadonnée est une donnée qui caractérise une autre donnée.

Pour un document numérique, les métadonnées sont contenues dans le fichier du document. Par exemple dans les propriétés des documents bureautiques sont enregistrés l'auteur, la dernière date de modification...

Pour une photo, les métadonnées seront la date, l'heure, la géolocalisation, etc...

Pour un fichier musique, les métadonnées seront le nom de l'artiste, le titre de la chanson, la photo de l'album, etc...

2° L'utilisation des métadonnées

Les principales utilisations des métadonnées seront :

- le classement des fichiers,
- la visualisation d'informations concernant l'origine du fichier,
- l'exploitation par les moteurs de recherche.

3° Vers un web sémantique

Le web actuel est un web de documents lisibles par des humains qui y interprètent les données pour produire des informations.

Les logiciels savent répertorier les mots mais ne leur donnent pas de sens.

Le web sémantique serait un web de données, à savoir que les données composant les documents seraient répertoriées pour leur sens et pourraient faire l'objet de traitements automatiques faisant émerger de nouvelles informations.

Le web sémantique ne consiste pas à créer de nouvelles données (elles sont déjà présentes) mais d'aider les humains à créer de nouvelles connaissances en interconnectant ces données

Question 5	La numérisation suffit-elle à valoriser l'information ?	
Chapitre 5.4	<i>Les métadonnées</i>	Bac STMG SIG
Cours		Lycée J. Feyder

4° Le stockage des métadonnées

Les principaux formats de stockage des métadonnées sont :

- le HTML qui propose plusieurs solutions (complémentaires):
 - ✓ les balises <meta> de la section <head> (toutefois Google ne tient plus compte des métadonnées keywords trop soumises à des manipulations),
 - ✓ les attributs rel des balises <a>, <ares> et <link> ,
 - ✓ les balises (HTML5) structurantes de pages (<header>, <nav>, <section>, <article>, <aside>, <footer>) ou de texte (<mark>, ...),
- le XML (ainsi un fichier Word .docx compresse différents fichiers, dont XML décrivant le document),
- le langage RDF (Resource Description Framework).