

Séance 4. Unicode

SNT - Thème 1. Codage de l'information

Unicode

Unicode

Unicode est un standard informatique conçu pour permettre le codage, la représentation et la gestion cohérente de texte exprimé dans la plupart des systèmes d'écriture du monde. Son objectif principal est de fournir un numéro unique pour chaque caractère, indépendamment de la plateforme, du programme ou de la langue.

Différences Majeures avec ASCII et ISO-Latin

ASCII (American Standard Code for Information Interchange)

- ▶ **Taille** : ASCII utilise 7 bits pour représenter chaque caractère, permettant ainsi de coder 128 caractères (de 0 à 127).
- ▶ **Jeu de caractères** : Inclut les lettres majuscules et minuscules de l'alphabet anglais, les chiffres, les signes de

Exercices

Exercice 1

Construire une frise chronologique des dates clés d'apparition et de standardisation de ces différents formats (ASCII, ISO-Latin, UNICODE)

Exercice 2

On dispose d'un fichier texte de 700 octets mais on ne sait pas quel encodage il utilise parmi ceux cités plus hauts. Quelle est la longueur du texte ...

1. S'il est encodé en ASCII sur 7 bits ?
2. S'il est encodé en ASCII sur 8 bits ?
3. S'il est encodé en iso-latin1 sur 8 bits ?
4. S'il est encodé en ASCII sur 8 bits ?