

Anecdote :

Le 18 Février 2021, l'astromobile *Perseverance* s'est posé avec succès sur Mars.
Son parachute était composé de parties rouges et blanches, symbolisant respectivement des 1 et des 0.

Le message "DARE MIGHTY THINGS", slogan du laboratoire JPL, était ainsi codé en binaire.

Repères historiques :

1854 : George Boole

Logicien britannique, introduit le calcul booléen qui se révélera très utile en informatique.

1963 : Naissance du code ASCII

Afin de représenter les caractères alphanumériques avec un ordinateur, le **code ASCII** (American Standard Code for Information Interchange) est développé pour permettre l'uniformisation du codage et faciliter les échanges d'informations.

A ses débuts, le code ASCII permettait uniquement de représenter les caractères principaux de la langue anglaise, ainsi que ses caractères "non imprimables" comme le retour à la ligne, codé sur 7 bits.

1986 : Apparition de la norme ISO 8859-1

Expression de la norme américaine ASCII, la norme **ISO 8859-1**, encore appelée Latin-1, est formée d'un ensemble de 191 caractères de l'alphabet latin, tous codés sur un unique octet.

Elle permet de représenter la très grande majorité des caractères de nombreuses langues d'Europe occidentale.

Des extensions existent pour d'autres alphabets, comme le cyrillique ou le polonais.

1988 : Le standard Unicode

Grâce au standard Unicode, les caractères de tous les systèmes d'écriture du monde peuvent posséder un nom et un identifiant numérique.

Les caractères représentés selon la norme U+xxxx où xxxx est un nombre composé de 4 à 6 caractères hexadécimaux.

L'**interopérabilité** entre différents logiciels devient possible.

Ce standard code actuellement 137 929 caractères.

1996 : La norme UTF-8 est adoptée

La **norme d'encodage UTF-8** est une représentation d'Unicode qui possède un encodage de taille variable selon le caractère à coder, ce qui permet de limiter le coût en mémoire et d'assurer la compatibilité avec la norme ASCII, plus ancienne.

Les caractères de la table ASCII restent ainsi codés sur 7 bits, tandis que les autres caractères sont représentés sur 2, 3 ou 4 octets.

L'encodage UTF-8 est utilisé par plus de 95% des sites web, en octobre 2020.

1. Combien de caractères différents peut-on coder sur 7 bits ?
2. Pouvez-vous imaginer une des limites du code ASCII ?
3. Quel est l'intérêt d'avoir ajouté un 8° bit de codage dans la norme Latin-1, en terme de quantité de caractères représentables ?
4. Dans le traitement de texte de votre choix, saisir Alt+65. Quel caractère de la table UTF-8 obtenez-vous ? Justifier que la caractère A soit codé par la valeur 41 en UTF-8.
5. A quel caractères de la table UTF-8 correspondent les saisies Alt+0192, Alt+0201 et Alt+0200 ?