

L'UTF-8 ne prend pas deux fois plus de place que l'encodage ISO-8859-15

culture informatique



L'encodage ISO-8859-15 utilise 1 octet (8 bits) pour représenter chaque caractère. Cela signifie qu'il peut représenter 256 caractères différents en utilisant un nombre limité d'octets. L'ISO-8859-15 est conçu pour couvrir les caractères utilisés dans les langues européennes occidentales, y compris le français, l'anglais, l'allemand et d'autres. Par conséquent, pour de nombreux textes dans ces langues, l'ISO-8859-15 peut suffire à représenter le contenu avec une taille raisonnable. Cependant, l'ISO-8859-15 n'est pas capable de représenter les caractères utilisés dans d'autres langues et peut ne pas être suffisant pour les applications qui nécessitent une plus grande variété de caractères.

L'UTF-8 peut utiliser de 1 à 4 octets (8 bits chacun) pour représenter un seul caractère. Cela signifie qu'il peut représenter plus de 1 million de caractères différents, ce qui en fait un encodage très flexible pour les applications qui nécessitent une variété de caractères pour différentes langues et scripts.

L'UTF-8 utilise une "variable longueur" de codage, ce qui signifie qu'il peut utiliser autant d'octets pour représenter les caractères les plus courants, tels que ceux utilisés en français et en anglais. En conséquence, pour certains textes, l'UTF-8 peut prendre autant de place que d'autres encodages, tels que l'ISO-8859-15, mais pas plus. Cependant, pour d'autres textes qui utilisent un grand nombre de caractères non-ASCII, l'UTF-8 peut nécessiter plus d'espace pour les représenter.

Non, l'UTF-8 ne prend pas deux fois plus de place que l'encodage ISO-8859-15 de manière générale. En fait, l'UTF-8 peut prendre plus ou moins de place que l'ISO-8859-15, selon la complexité des caractères codés. L'UTF-8 est conçu pour prendre autant de place que l'ISO-8859-15 pour les caractères les plus courants, mais peut prendre plus de place pour les caractères plus rares ou complexes.

L'UTF-8 peut utiliser jusqu'à 4 octets pour représenter un seul caractère, ce qui est plus que le 1 octet (8 bits) utilisé par l'ISO-8859-15. Cependant, l'UTF-8 est conçu pour être aussi compact que l'ISO-8859-15 pour les caractères courants, en utilisant autant d'octets pour les représenter. L'UTF-8 utilise une variable longueur de codage pour représenter les caractères. De plus, l'UTF-8 est conçu pour être compatible avec les codes ASCII, qui sont utilisés pour représenter les caractères les plus courants dans de nombreuses langues, y compris le français et l'anglais. En conséquence, pour de nombreux textes en français ou en anglais, l'UTF-8 peut effectivement prendre autant de place que l'ISO-8859-15, mais pas plus.



— [Cédric ABONNEL \(cedricabonnel\)](#), *CPT & Wikipedia* - Article paru initialement le Lundi 6 Février 2023

From:
<https://www.abonnel.fr/> - notes informatique & technologie

Permanent link:
https://www.abonnel.fr/journal_geek/2023/20230206-l-utf-8-ne-prend-pas-deux-fois-plus-de-place-que-l-encodage-iso-8859-15

Last update: **2023/02/06 21:43**

