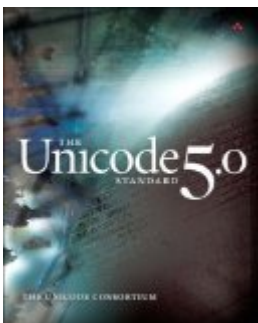## Difference Between Similar Terms and Objects

Custom Search   Search

- **Business**
  - **Accounting**
  - **Economics**
  - **Finance**
  - **Investment**
  - **Management**
  - **Marketing**
  - **Organizations**
  - **Planning & Activities**
  - **Product & Services**
  - **Structure & Systems**
- **Miscellaneous**
  - **Career & Education**
  - **Culture**
  - **Entertainment**
  - **Fashion & Beauty**
  - **Geography**
  - **History**
  - **Legal**
  - **Politics**
    - **Ideology**
    - **Leaders**
    - **Political Institutions**
  - **Religion**
    - **Islam**
  - **Sports**
- **Technology**
  - **Gadgets**
  - **Gaming**
  - **SmartPhones**
  - **Tablets**
  - **Software**
  - **internet**
  - **Hardware**
  - **Protocols & Formats**
  - **Communication**
  - **Web Applications**

Difference Between Unicode and UTF-8

• Categorized under **Protocols & Formats**,**Technology** | **Difference Between Unicode and UTF-8**

Unicode vs UTF-8

The development of Unicode was aimed at creating a new standard for mapping the characters in a great majority of languages that are being used today, along with other characters that are not that essential but might be necessary for creating the text. UTF-8 is only one of the many ways that you can encode the files because there are many ways you can encode the characters inside a file into Unicode.

UTF-8 was developed with compatibility in mind. ASCII was a very prominent standard and people who already had their files in the ASCII standard might hesitate in adopting Unicode because it would break their current systems. UTF-8 eliminated this problem as any file encoded that only has

characters in the ASCII character set would result in an identical file, as if it was encoded with ASCII. This allowed people to adopt Unicode without needing to convert their files or even changing their current legacy software that was unaware of the Unicode standard. Any of the other mapping methods for Unicode breaks compatibility with ASCII and would force people to convert their system.

The observance of compatibility to ASCII of **UTF-8** produces a side-effect that makes it ideal for word processing where most of the time, all the characters being used are included in the ASCII character set. UTF-8 only uses a byte to represent every code point resulting in a file size that is half to the same file encoded in UT-16 which uses 2 bytes, and a quarter to the same file encoded in UTF-32 which uses 4.

UTF-8 has been adopted in the World Wide Web because it is both space efficient and byte oriented. Web pages are often simple text files that usually do not contain any character that is outside the ASCII character set. Using other **encoding methods would only increase** the network load without any benefit. Even in email transport systems, UTF-8 is slowly but surely being adopted as a replacement for the older encoding systems that are still being used.

Summary:

1. Unicode is the standard for computers to display and manipulate text while UTF-8 is one of the many mapping methods for Unicode
2. UTF-8 is a mapping method the retains compatibility with the older ASCII
3. UTF-8 is the most space efficient mapping method for Unicode compared to other encoding methods
4. UTF-8 is the most used Unicode standard for the web

Author	Recent Posts

Ben Joan

| 2 | 0 |

Search DifferenceBetween.net :

[   ] [Search] Google   Custom Search

⭐⭐⭐⭐⭐Help us improve. Rate this post! (**9** votes, average: **4.78** out of 5)

✉️   **Email This Post** : If you like this article or our site. Please spread the word. Share it with your friends/family.

- **Difference Between UTF-8 and UTF-16**
- **Difference Between Unicode and ASCII**
- **Difference Between ANSI and ASCII**
- **Difference Between ANSI and UTF-8**
- **Difference Between EBCDIC and ASCII**

Cite
Ben Joan. "Difference Between Unicode and UTF-8." DifferenceBetween.net. June 22, 2017 < http://www.differencebetween.net/technology/difference-between-unicode-and-utf-8/ >.

Leave a Response

Written by : Ben Joan. and updated on June 22, 2017

See more about : **character**, **unicode**