**ESCUELA POLITÉCNICA NACIONAL**

ESCUELA DE FORMACIÓN DE TECNÓLOGOS

**SISTEMAS OPERATIVOS II**

|  |  |
| --- | --- |
|  | |
| ASIGNATURA: | Sistemas Operativos II |
| PROFESOR: | Ing. Juan Zaldumbide |
| PERÍODO ACADÉMICO: | Abril. 2016 - Sep. 2016 |
|  | |

**CONSULTA 1**

|  |  |
| --- | --- |
|  | |
| TÍTULO:  **UNICODE UTF-8** | |
| **ESTUDIANTE** | |
| Rubén Pozo | |
|  | |
|  | |
|  | |
|  | |
|  | |
| FECHA DE REALIZACIÓN: | 19 de Mayo de 2016 |
| FECHA DE ENTREGA: | 26 de Mayo de 2016 |
| CALIFICACIÓN OBTENIDA: |  |
| FIRMA DEL PROFESOR: |  |
|  | |

# CONTENIDO

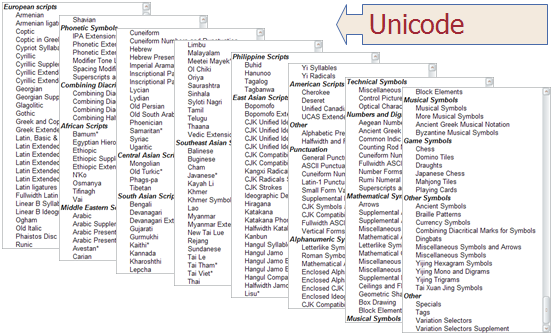
**Unicode:**

Es un set de caracteres universal o estándar para la escritura en la mayoría de los idiomas hablados que se usan en la computadora actualmente. Su principal objetivo es ser un supe conjunto de todos los sets de caracteres que se hayan codificado. El texto que nosotros vemos en la web o en cualquier editor de texto como letras, signos de puntuación, números, símbolos por detrás está representado mediante caracteres.

Organizaciones en el pasado, han recopilados diferentes sets de caracteres y han logrado crear codificaciones especiales para ellos. Un set solo puede abarcar idiomas de Europa con base en el latín sin incluir países como Bulgaria o Grecia que cuentan con otro tipo de caracteres especiales. Otro set puede contemplar un idioma específico como el japonés, y otros pueden ser parte de distintos sets diseñados especialmente para representar algún idioma en específico.

Lamentablemente, no es posible garantizar que su aplicación particular pueda soportar todas las codificaciones, ni que una determinada codificación pueda soportar todos sus requerimientos para la representación de un cierto idioma. Además, generalmente resulta imposible combinar distintas codificaciones en la misma página web o en una base de datos, por lo que siempre es muy difícil soportar páginas plurilingües si se aplican enfoques "antiguos" cuando se trata de tareas de codificación.

Con este sistema Unicode, podemos abarcar todos los caracteres necesarios para cualquier sistema de escritura en el mundo, incluyendo lenguajes tan antiguos como el cuneiforme, gótico y jeroglíficos egipcios. Este enfoque facilita mucho el trabajo con sistemas de páginas en varios idiomas y se adapta mejor a las necesidades de los usuarios.



**UTF-8:**

Que por sus siglas en inglés quiere decir (8 bit Unicode Transformation Format), Este formato utiliza símbolos de longitud variable y fue definido como estándar de codificación reconocida por Unicode y lenguajes web

Este formato puede representar cualquier carácter Unicode, usa símbolos de 1 a 4 bits por carácter Unicode, Incluye especificación ASCII de 7 bits, es posible determinar el inicio de cada símbolo sin reiniciar la lectura desde el principio de la comunicación, no se superpone por lo cual no es posible confundirlos entre sí. Todas estas características lo hacen el atractivo para adoptarlas en la codificación de correo electrónico y páginas web.

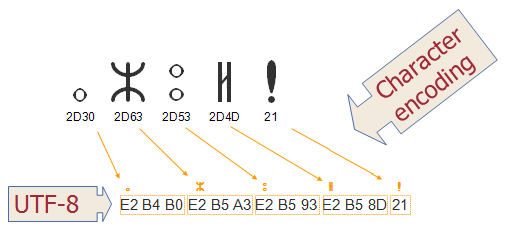
UTF-8 permite codificar utilizando tamaños de datos de 8 bits y funciona bien con muchos de los sistemas operativos existentes.

Las referencias numéricas de caracteres especifican la posición del código de un carácter en el conjunto de caracteres del documento. Las referencias numéricas de caracteres pueden tener dos formas:

* La sintaxis "&#D;", donde D es un número decimal
* La sintaxis "&#xH;" o "&#XH;", donde H es un número hexadecimal (Para los números hexadecimales de referencias de caracteres numéricas no se distingue entre mayúsculas y minúsculas.)

Para el intervalo ASCII de caracteres, UTF-8 es idéntico a la codificación ASCII y permite un conjunto mayor de caracteres. Para scripts español y griego, sin embargo, UTF-8 puede necesitar dos bytes para cada carácter.

|  |  |  |
| --- | --- | --- |
| **Descripcion** | **Caracter** | **Código** |
| Left angle quotes | « | &#171; |
| Lower-than sign | < | &#60; |
| Plus-minus sign | ± | &#177; |
| Multiplication sign | × | &#215; |
| not equal to | ≠ | &#8800; |
| integral | ∫ | &#8747; |
| fraction one-quarter | ¼ | &#188; |
| Euro | € | &#8364; |
| Franc | ₣ | &#8355; |
| Peseta | ₧ | &#8359; |
| Capital O with slash | Ø | &#216; |
| Capital Y with acute | Ý | &#221; |



# CONCLUSIÓNES

En conclusión el formato Unicode UTF-8 nos sirve para representar todos los caracteres que nosotros consideramos como letras o símbolos y en especial caracteres especiales como la “ñ”, de esta manera nos amplía la posibilidad de expresar màs símbolos o caracteres dentro de nuestra escritura por computadora.

# RECOMENDACIONES

Es recomendable aprender cómo funciona este formato Unicode ya que la mayor parte de páginas web y correos electrónicos se manejan bajo este estándar.

# BIBLIOGRAFÍA

# Bibliografía

|  |  |
| --- | --- |
| [1] | R. Ishida, «Internacionalización,» W3C, 20 Octubre 2010. [En línea]. Available: https://www.w3.org/International/articles/definitions-characters/index.es. [Último acceso: 25 Mayo 2016]. |
| [2] | «Periodni,» Eni Generalic, 1 Sptiembre 2006. [En línea]. Available: http://www.periodni.com/es/codificacion\_utf-8\_unicode.html. [Último acceso: 25 Mayo 2016]. |