

Perl:
WWW
HTML
CGI

WWW

- HTML es llamado el lenguaje del World Wide Web o simplemente WWW o “web”
 - El web comenzó en 1989 con Tim Berners-Lee quien investigaba las formas de compartir información
 - En 1992 se hace público un protocolo e interfaz, el cual se adoptó rápidamente y nació el web
- La mayoría de los estándares del web son definidos por medios de RFCs, los cuales se pueden obtener desde:
 - <http://www.ietf.org>

WWW

- Documentación del web se encuentra en “World Wide Web Consortium”:
 - <http://www.w3.org>
- Terminología:
 - HTTP: hypertext transfer protocol, protocolo para transferir información entre un cliente y un computador servidor
 - MIME: Multimedia Internet Mail Extensions, comenzó como una extensión del email, para permitir el intercambio de información distinto a texto plano (mas info RFC 1521)
 - SGML: Standard Generalized Markup Language, estándar internacional para definir documentos estructurados

WWW

- Terminología:
 - HTML: Hypertext Markup Language, aplicación particular de SGML, el cual define el lenguaje usado en WWW
 - XML: Extensible Markup Language, derivado de SGML, mas simplificado, se puede personalizar de acuerdo a las necesidades
 - CSS: Cascading Style Sheets, estándar internacional que permite separar los contenidos de “estilo” de un documento HTML
 - URL: Uniform Resource Locator, forma de especificar la ubicación de un recurso, consiste en un protocolo, un “:” y dos “/”, nombre de máquina y path del recurso

HTML

CC31A Programación SW de Sistemas

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0  
Transitional//EN"  
"http://www.w3.org/TR/REC-html40/loose.dtd">  
<HTML>  
<HEAD>  
<TITLE>Simple HTML Document</TITLE>  
</HEAD>  
<BODY>  
<H1>Título</H1>  
<P>Texto de ejemplo</P>  
</BODY>  
</HTML>
```

José Urzúa R jourzua@dcc.uchile.cl

HTML

- CC31A Programación SW de Sistemas
- Las partes del documento que se encuentran entre “<>” no son desplegadas
 - Estas controlan el despliegue del texto restante
 - La sección “<!DOCTYPE ...” define que es un documento que cumple el estándar HTML 4.0 “transitional”
 - Si el documento no cumple un estándar, se debería omitir esta sección
 - Existen características de algunos software que parecen muy buenas, pero no respetan algún estándar
 - Para no tener impedimentos de visualización de una página HTML, es ideal utilizar un estándar

José Urzúa R jourzua@dcc.uchile.cl

HTML

- CC31A Programación SW de Sistemas
- El resto del documento es encerrado en los tags “<HTML></HTML>” los cuales definen el inicio y fin del código HTML
 - Luego hay dos secciones: head y body
 - Head: principalmente se coloca el título de la página el cual generalmente es desplegado en la parte superior del browser
 - Dentro del body hay dos partes, la delimitada por <H1> que indica un “header” de nivel 1
 - Otra parte es <P> que indica un párrafo

José Urzúa R jourzua@dcc.uchile.cl

HTML

- CC31A Programación SW de Sistemas
- Los tags no son case-sensitive:
 - Es lo mismo <TITLE>, <Title>, <title>
 - Para el título del documento es útil colocar palabras descriptivas del texto, para poder reconocerlo dentro de un bookmark por ejemplo
 - Para agregar un comentario dentro de un documento HTML:

 <!-- este es un comentario -->

José Urzúa R jourzua@dcc.uchile.cl

HTML

- Para destacar títulos dentro del texto existe 6 tipos de “headings” desde H1 hasta H6
- El formateo del texto se realiza de manera automática
 - Es bueno por que el autor no sabe el ancho de la ventana que visualizará la página
 - Con
 se indica un salto de línea
 - <P> </P> Indica que un nuevo párrafo comenzará
 - <HR> dibuja una línea horizontal
- Estilos de texto:
 - <ADDRESS></ADDRESS>
 - Texto en negrita

HTML

- Estilos de texto:
 - <I></I> italic
 - <PRE></PRE> texto pre-formateado
 - subscript
 - superscript
- Caracteres especiales:
 - < < ´ á
 - > > æ æ
 - & & ¸ ç
 - " “ ü ü
 - © ©
 - (format.html)

HTML

- Para desplegar listas:
 - No ordenadas:
 - Directorio: <DIR> </DIR>
 - Menú: <MENU> </MENU>
 - (list.html)
- Imágenes:
 - Para agregar imágenes al texto basta con utilizar el tag
 - Con el atributo “src” se indica la ubicación de la imagen, podría ser un archivo o una URL
 - JPEG, GIF y PNG son utilizados comúnmente
 - (beer.html)

HTML

- Tablas:
 - Útiles para formatear datos
 - Se utilizan los tags <TABLE></TABLE>
- Solo “dentro” de la definición de una tabla son válidos:
 - <CAPTION></CAPTION>
 - <TR></TR> fila
 - <TH></TH> header de la tabla
 - <TD></TD> celda
 - (table1.html)
 - (table2.html)

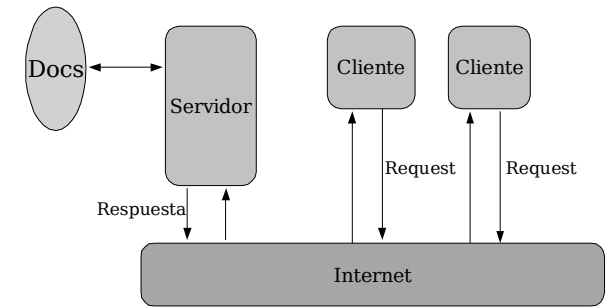
HTML

- Enlaces:

```
<a href="URL">click here!</a>
```

- URL puede ser otro sitio, un archivo o un enlace dentro del mismo archivo
- (link.html)
- (anchor3.html)

HTML



CGI

- “Common Gateway Interface” especificación que define como la información es pasada desde un browser a un servidor
 - El programa que recibe dicha información (programa CGI) procesa esa información y usa el protocolo HTTP para enviar el documento dinámico al browser
- Por medio de los tags <FORM> se puede recibir información desde los clientes
 - Se debe especificar el atributo “action”, que indica la URL del programa que se invocará para procesar el formulario
 - Se puede especificar el método por medio del cual se enviarán los parámetros al CGI

CGI

- Para solicitar información al usuario, se utilizan los INPUT tags, que tienen varios tipos:
 - texto: el browser muestra un cuadro simple en el cual se puede escribir texto
 - password: igual al anterior, pero el texto que se escribe no se despliega
 - hidden: campo de texto que no se despliega en el browser, y el valor es asignado por el programa del servidor
 - checkbox: permite al usuario seleccionar una o varias opciones de las desplegadas
 - radio: igual al checkbox, excepto que sólo se puede seleccionar una opción

CGI

- INPUT tag:
 - submit: despliega un botón en el browser, cuando se presiona, se ejecuta el programa indicado en el "action" del formulario, enviando todos los valores
 - reset: despliega un boton en el browser, al presionarlo cambia todos los valores desplegados a los iniciales (cuando se cargo el formulario por primera vez)
 - select: permite seleccionar desde una lista de valores
 - textarea: despliega una caja de texto de múltiples líneas y columnas

CGI

- Para escribir programas CGI en Perl, existe el módulo CGI.pm
- Para que el programador escriba menos código HTML, este módulo implementa variadas funciones útiles
- Por ejemplo:
 - header(): retorna un string con el "Content-type"
 - start_html(string): retorna "string" como título del documento html
 - h1(string): retorna <H1>string</H1>
 - p(string): retorna "string" como un nuevo párrafo

CGI

- Para utilizar estas funciones se deben importar por medio del uso de tags:
 - :cgi importa los métodos de manipulación de argumentos
 - :form importa todos los métodos para generar código HTML de un form, por ejemplo: textfield()
 - :standard importa todos los métodos estándares, generalmente se utiliza este flag
- (ejemplo.pl)
- (ejemplo2.pl)
- Ejemplo de formulario: cgi1.htm y procesa.pl