

## Tema 2

# Lenguaje de marcas HTML (I)

### Objetivos

- Identificar y clasificar los lenguajes de marcas relacionados con la web y sus diferentes versiones.
- Analizar la estructura de un documento HTML.
- Identificar las partes que componen un documento HTML.
- Identificar las semejanzas y diferencias entre los lenguajes HTML y XHTML.
- Reconocer la utilidad del lenguaje HTML en los sistemas de gestión de la información.
- Conocer la funcionalidad de las principales etiquetas y atributos del lenguaje HTML.

### Introducción

En el tema anterior se introdujeron los conceptos relacionados con los lenguajes de marcas, así como los conceptos básicos que permiten comprender el contexto en el que se aplican. En este tema se inicia el estudio del lenguaje de marcas HTML en relación con la construcción de páginas y sitios Web. El estudio del lenguaje HTML se ha estructurado en varios temas debido a su amplitud, facilitándose de este modo su aprendizaje de una forma más detallada, ordenada y concreta. Estos temas se han planificado de forma consecutiva dentro de la periodificación de contenidos del módulo profesional.

### Índice

2.1 ASPECTOS PREVIOS: EL SERVICIO DE INFORMACIÓN WEB	2
2.2 HTML: UN LENGUAJE PARA LA WEB	3
2.3 DOCUMENTOS HTML	5
2.4 ESTRUCTURA DE UN DOCUMENTO HTML	7
2.5 ELEMENTOS BÁSICOS	12
2.6 FORMATO DEL TEXTO	18
2.7 VALIDACIÓN DEL CÓDIGO HTML	22
Glosario de términos	23

## 2.1 ASPECTOS PREVIOS: EL SERVICIO DE INFORMACIÓN WEB

El servicio de información Web, también denominado WWW o *World Wide Web*, representa un universo de información accesible a través de Internet. Está formado por un conjunto de recursos de información interconectados que conforman buena parte del conocimiento humano actual. Su funcionamiento es posible debido a la existencia de una serie de elementos lógicos (software) y físicos (hardware).

El servicio de información Web puede definirse como un **servicio hipermedia de acceso a la información a través de internet**, que está basado en:

- El uso de la infraestructura de Internet basada en la familia de protocolos TCP/IP.
- La utilización de documentos hipertexto para facilitar el acceso y la navegación entre los recursos solicitados mediante enlaces, también denominados hipervínculos o *links*.
- El modelo de arquitectura del software Cliente/Servidor que incorpora un esquema de petición-respuesta para el intercambio de información entre ambos.
- El sistema *Uniform Resource Identifier* (URI) para la identificación genérica de los recursos en la red. Un identificador URI referencia unívocamente un recurso de la red mediante una cadena de caracteres, empleando la siguiente sintaxis: servicio:autoridad/ruta/recurso. Por ejemplo, la URI: <ftp://ftp.iesmarenostrum.com/pub/Tema1.doc>, referencia al recurso Tema1.doc empleando el servicio ftp. En el caso concreto de referencias a direcciones Web, se suele emplear el sistema de identificación particular del servicio de información Web que se denomina *Uniform Resource Locator* (URL). Así, el concepto URL se ha incorporado al concepto general, que es el sistema URI. Un ejemplo de una URL, o de una URI para la Web, sería: <http://www.iesmarenostrum.com/informatica/pub/inicio.htm>.
- El protocolo *HyperText Transfer Protocol* (HTTP) como mecanismo de transmisión de los recursos solicitados por el cliente. Se trata del protocolo de comunicaciones usado para el intercambio de información en cada transacción entre el cliente y el servidor en el servicio de información Web.

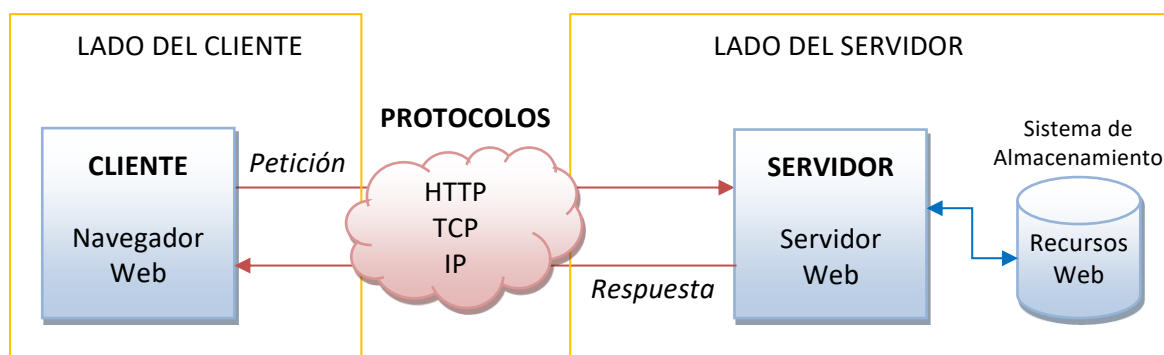
El término hipermedia que aparece en la definición anterior, relativa al servicio de información Web, debe entenderse en el sentido de multimedia. Es decir, como un servicio de información que puede presentar información procedente de medios de comunicación diversos: texto, gráficos, sonido, videos, etc. Actualmente, el uso del lenguaje se simplifica y suele ser habitual referirse al servicio de información Web, sencillamente, como: la Web.

### Funcionamiento del servicio de información Web

El mecanismo de funcionamiento del servicio de información Web es muy sencillo. Se basa en un modelo de arquitectura de software Cliente/Servidor, en el cual el intercambio de información entre el cliente y el servidor se realiza mediante un sistema de petición-respuesta.

A continuación, se describe detalladamente el funcionamiento del servicio de información Web:

1. En primer lugar, el programa navegador, o cliente Web, utiliza una URL, para conectarse con un programa servidor, o servidor Web, y solicitarle un recurso que contiene la información multimedia solicitada. Este recurso será un documento HTML.
2. Establecida la conexión con el servidor Web, en segundo lugar, éste localiza el documento HTML dentro de su sistema de almacenamiento asociado y lo envía al cliente como respuesta a la solicitud realizada previamente, a través de la infraestructura de Internet.
3. Finalmente, conforme va llegando al cliente el contenido del documento HTML, éste va siendo interpretado por el programa navegador o cliente Web, reconstruyéndose la información solicitada en la ventana de visualización.



La acción de publicar o desplegar un recurso, como puede ser un documento HTML, en un servidor Web consiste, básicamente, en ponerlo a disposición de los potenciales visitantes a través del servicio Web de Internet. Para ello, basta con almacenar el recurso a compartir en una carpeta de publicación del sistema de almacenamiento de un servidor Web activo.

Como puede comprobarse, el papel del lenguaje HTML es fundamental en el funcionamiento del servicio de información Web. Permite construir documentos destinados a la presentación de información a través del servicio de información Web de Internet. Además, al tratarse de texto plano, los documentos HTML pueden ser interpretados por cualquier navegador y, por tanto, pueden ser procesados utilizando cualquier plataforma, dispositivo, sistema operativo, etc. La aplicación más significativa del lenguaje HTML es la creación de interfaces de usuario para Web, con las cuales los usuarios del servicio de información Web pueden interactuar. Sin embargo, el campo de aplicación de HTML es más amplio que su uso en el servicio de información Web. También puede utilizarse para construir sistemas de ayuda de ejecución local, distribución de contenidos sobre diversos tipos de soportes de información, etc.

## 2.2 HTML: UN LENGUAJE PARA LA WEB

Cuando se visita una página Web puede apreciarse que está formada por textos, imágenes, colores de fondo, en algunos casos elementos que se desplazan por la pantalla, vídeos, etc. en fin, todo el conjunto de elementos que hacen de dicha página un espacio multimedia e interactivo. Aunque se

podría pensar que todos esos elementos están incluidos en la página, no es así. Realmente, el único de todos ellos que pertenece a la propia página es el texto. Y, entonces, ¿por qué se ve todo lo que aparece en la página? La respuesta está en comprender que una página web no es más que un documento de texto acompañado de una colección de etiquetas que sirven para que el navegador Web: interprete de qué forma tiene que presentar el texto; sepa de dónde tiene que obtener las imágenes, los sonidos y el resto de los recursos de la página, así como de qué manera mostrarlos; y, distribuya todos los contenidos de una forma predeterminada en la ventana de visualización del navegador. El lenguaje de marcas HTML permite definir todas estas funcionalidades, empleando para ello un conjunto de etiquetas y atributos definidos en el lenguaje.

Las siglas HTML (*Hyper-text Markup Language*) se corresponden con la traducción "Lenguaje para marcado de hipertexto". Este lenguaje de marcas, por tanto, se utiliza para estructurar documentos de texto que puedan ser interpretados por navegadores y distribuidos a través del servicio de información Web de Internet.

### Evolución histórica del lenguaje HTML

La primera versión del lenguaje HTML fue desarrollada por Tim Berners-Lee en 1989 en el CERN, el Consejo Europeo para la Investigación Nuclear con sede en Suiza. Durante los siguientes años el lenguaje fue ampliando sus especificaciones, al ritmo al que se fue ampliando su difusión y aceptación. El éxito alcanzado por el lenguaje HTML y del servicio de información Web se ha basado en su sencillez y en el uso a nivel general, prácticamente, de las mismas especificaciones a lo largo del tiempo, lo que ha motivado el esfuerzo de los usuarios y desarrolladores por asegurar esa especificación. La evolución de HTML como lenguaje formal se ha desarrollado ampliando sus funcionalidades, pero manteniendo básicamente su especificación básica.

- En 1994 se publica el primer estándar oficial llamado HTML 2.0.
- En 1995 se publica la versión HTML 3.0 que no obtuvo un consenso generalizado entre la comunidad de usuarios y desarrolladores.
- En 1997 se publica la versión HTML 3.2 por parte del consorcio W3C (*World Wide Web Consortium*), que es el organismo oficial internacional e independiente encargado de publicar los estándares del lenguaje.
- En 1998 aparece la versión HTML 4.0 que incluye soporte para lenguajes script, hojas de estilo, mejoras en tablas y formularios, facilidades de impresión, uso de distintos formatos de escritura y alfabetos, incrustación de objetos multimedia, etc.
- En 1999 aparece **HTML 4.01** que incorpora mejoras en cuanto a la accesibilidad Web.
- En 2000 el consorcio W3C publica la especificación **XHTML 1.0** que es una reformulación de HTML 4.0 basada en las reglas del lenguaje XML. Esta especificación está extensamente aceptada y se utiliza ampliamente.

- En 2004 los fabricantes de navegadores Apple, Mozilla y Opera crean el grupo de trabajo WHATWG (*Web Hypertext Application Technology Working Group*) para seguir con el desarrollo del lenguaje de marcas HTML.
- En 2008 W3C y WHATWG, unidos, publican el borrador de la especificación **HTML5** que incluye todos los elementos válidos tanto en HTML 4.0 como en XHTML 1.0 y, además, incorpora novedades interesantes. Funciona en la mayoría de los navegadores existentes actualmente y, además, es compatible con los navegadores antiguos.

En la actualidad, el código HTML se suele escribir empleando las especificaciones HTML 4.0 y XHTML 1.0 que son las especificaciones más utilizadas. Sin embargo, la especificación HTML 5 va ganando terreno progresivamente, debido a las mejoras que incorporan sus nuevas funcionalidades.

## 2.3 DOCUMENTOS HTML

Un documento HTML es un **archivo de texto plano** en formato ASCII que incluye una serie de **etiquetas** o marcas que definen cada uno de los elementos que lo componen. A su vez, las etiquetas pueden incluir **atributos** para definir sus características. Los documentos HTML se almacenan en **archivos cuyo nombre tienen la extensión .html**. Los documentos HTML constituyen las páginas Web que están accesibles a través del servicio de información Web.

Para crear un documento HTML es necesario utilizar un programa editor de texto, para escribir el código, y un programa navegador Web, para que interprete el código HTML y se pueda comprobar el resultado obtenido. Dado que los documentos HTML están formados por texto plano, puede utilizarse cualquier editor de texto convencional para crearlos y editarlos. Se propone utilizar el editor de código **Notepad++**, aunque pueden emplearse otros como: UltraEdit, Brackets, Sublime Text, Visual Studio Code, etc. En cuanto al navegador, se recomienda utilizar varios de ellos, especialmente los más extendidos, para visualizar la presentación de los documentos HTML que se van creando. De este modo, es posible comprobar el comportamiento de todos los elementos que forman los documentos HTML que se crean y, además, comparar las diferencias existentes en cuanto a la interpretación que realizan del código cada uno de ellos. Los navegadores más extendidos actualmente pueden ser: Microsoft Edge, Google Chrome, Mozilla Firefox, Safari, etc.

En principio, no será necesario publicar los documentos HTML en un servidor Web de Internet. Los navegadores pueden interpretar archivos locales que incluyen código HTML. Es decir, pueden interpretar el código HTML contenido en los archivos que están almacenados en el sistema de archivos del ordenador local, sin necesidad de tener que publicarlos en un servidor Web.

### Etiquetas y atributos en documentos HTML

Las **etiquetas** de HTML se escriben entre los símbolos de menor que (<) y de mayor que (>). Normalmente, se utilizan dos etiquetas: una de inicio o apertura y otra de fin o cierre. La única diferencia entre ambas es que la de cierre incluye una barra inclinada (/) antes del nombre. De esta manera, el texto que aparezca entre las etiquetas de inicio y de fin constituye el **contenido de información** del documento que se presentará con el efecto que dichas etiquetas indiquen.

En el siguiente ejemplo, la etiqueta <b> establece el texto en negrita, la etiqueta <i> en cursiva y la etiqueta <font> se utiliza, en este caso, para especificar un tipo de fuente Arial.

```
<b>Este texto aparecerá en negrita en la pantalla.</b>  
<i>Este texto aparecerá en cursiva en la pantalla.</i>  
<font family="Arial">Este texto se mostrará en con un tipo de fuente Arial.</font>
```

En los siguientes apartados, se irá estudiando cada una de las etiquetas con mayor profundidad.

Se pueden escribir etiquetas dentro de otras etiquetas. Sin embargo, en estos casos hay que tener en cuenta que se deben cerrar en orden inverso al orden de apertura, siguiendo la sintaxis general que se indica a continuación:

```
<etiqueta1>  
  Este nivel está afectado por la etiqueta 1  
  <etiqueta2>  
    Este nivel está afectado por la etiqueta 1 y 2  
    <etiqueta3>  
      Este nivel está afectado por la etiqueta 1, 2 y 3  
    </etiqueta3>  
    Este nivel está afectado por la etiqueta 1 y 2  
  </etiqueta2>  
  Este nivel está afectado por la etiqueta 1  
</etiqueta1>
```

Como puede apreciarse, para facilitar esta tarea se suelen dejar sangrías en el código cada vez que se incluye una etiqueta de apertura de un nuevo elemento. Por ejemplo:

```
<p>  
  texto normal  
  <b>  
    Texto en negrita  
    <i>  
      Texto en negrita y cursiva  
      <small>  
        Este texto aparecerá en negrita, cursiva y tamaño pequeño.  
      </small>  
      Texto en negrita y cursiva  
    </i>  
    Texto en negrita  
  </b>  
</p>
```

La etiqueta <p> marca el inicio de párrafo y la etiqueta <small> establece un tamaño de fuente pequeño.

La mayoría de las etiquetas admiten **atributos**. Los atributos especifican las características que tendrán las etiquetas sobre las que se aplican. La sintaxis general de las etiquetas con atributos es el siguiente:

```
<etiqueta1 atributo1="valor1" atributo2="valor2">  
  Texto afectado por la etiqueta1  
</etiqueta1>
```

Cada etiqueta tiene establecidos sus propios atributos disponibles. Los valores de los atributos pueden ser de tipo numérico, texto, etc., pero siempre deberán ir entre comillas. Por ejemplo, para establecer un tipo de fuente Arial se puede emplear el atributo family de la etiqueta <font>:

```
<font family="Arial">Este texto se mostrará en con un tipo de fuente Arial.</font>
```

Como ya se sabe, un documento HTML está formado, exclusivamente, por texto. Una parte de ese texto se presentará en la ventana de visualización del navegador, porque constituye el contenido de información que se desea compartir. El resto del texto del documento HTML está formado por etiquetas y atributos de HTML que no se presentarán en el navegador cuando se reconstruya la página, pero que, sin embargo, tendrán algún efecto sobre la presentación visual de la página. Estos efectos pueden ser, por ejemplo: poner parte del texto en negrita, estructurar los contenidos en forma de tabla, cambiarán el tipo de letra, aumentar o disminuir el tamaño del texto, facilitar al usuario el acceso a otros contenidos a través de enlaces, etc.

Las etiquetas, así como sus atributos, se escriben utilizando letras minúsculas. Además, a la hora de construir un documento HTML deben seguirse una serie de reglas en cuanto a la organización del documento, para asegurar que los navegadores Web puedan interpretar correctamente el código HTML incluido, de manera que se obtengan los resultados esperados. A continuación, se aborda cómo ha de ser la estructura de los documentos HTML.

## 2.4 ESTRUCTURA DE UN DOCUMENTO HTML

Los documentos HTML poseen una estructura compuesta, básicamente, por dos secciones: la cabecera y el cuerpo del documento. Estas secciones quedan definidas en el código mediante as etiquetas correspondientes. La sección de cabecera se utiliza para definir diversa información del documento HTML, como puede ser el título del documento HTML y otra información adicional, y, también para especificar diversos aspectos de configuración del propio documento y del navegador Web. Mientras que el cuerpo del documento contiene los distintos elementos que forman el documento: párrafos, títulos, listas, tablas, imágenes, etc. que se reconstruirán en la ventana de visualización del navegador cuando el navegador interprete el código HTML. La estructura general de un documento HTML tiene el siguiente formato:

### DECLARACIÓN DE TIPO DE DOCUMENTO

```
<html lang="xx">  
  <head>  
    <title>Título de la página </title>  
    Elementos opcionales de la cabecera  
    <style>Definición de estilos de la página</style>  
  </head>  
  <body>  
    Contenido del cuerpo  
  </body>  
</html>
```

El atributo lang de la etiqueta <html> define el idioma principal de la información contenida en el documento HTML, lo que facilita el trabajo a los motores de búsqueda y a los navegadores Web.



Como puede apreciarse, el contenido del documento HTML está estructurado en tres partes bien diferenciadas:

- **Declaración del tipo de documento.** Se trata de una única línea de texto que especifica el tipo de documento de qué se trata. De entre los tipos de documentos válidos, cada uno de ellos tiene unas características diferentes que pueden ajustarse mejor o peor a cada caso concreto. Los tipos de documentos se estudian a continuación, en este mismo apartado. Es la única parte del documento que no está incluida dentro de las etiquetas `<html>` `</html>`.
- **Cabecera.** Está delimitada por las etiquetas `<head>` `</head>`. En la cabecera se incluye la definición de diversa información relativa a la propia página Web que no va a presentarse en la ventana de visualización del navegador, salvo el título del documento que aparece en la barra superior del navegador web. Habitualmente, se utiliza para indicar el nombre del autor, las palabras clave destinadas a los buscadores, las hojas de estilo externas que se utilizarán con la página, etc. Además, también se utiliza para especificar diversos aspectos de configuración del propio documento o del navegador, como puede ser la codificación de caracteres que utiliza la página.
- **Cuerpo.** Está delimitado por las etiquetas `<body>` `</body>`. En el cuerpo de un documento HTML se incluye el contenido de información que se va a presentar la ventana de visualización del navegador, así como las etiquetas que definen la apariencia de la presentación de esos contenidos. Con la etiqueta `<body>` se delimita la parte de un documento HTML en la que se incluye el contenido en sí del documento. En el cuerpo es donde se especifican los elementos que conforman la página Web: el texto, las imágenes, los enlaces, las tablas, las listas, las divisiones, etc.

### La declaración del tipo de documento

El W3C es un organismo que se encarga de fijar las normas que deben seguir los documentos escritos utilizando el lenguaje HTML. Según este organismo, todo documento HTML deberá contener al principio del mismo una línea que indique qué tipo de documento es. Es lo que se llama la declaración del tipo de documento. Al indicar la declaración del tipo de documento o DTD (*Document Type Definition*) al principio del documento se está indicando al navegador cómo debe interpretar el código definido en el documento HTML. Si se emplean las especificaciones HTML 4.01 o XHTML 1.0, los tipos de documento válidos y sus declaraciones correspondientes son:

- **Documento HTML estricto.** Los documentos de este tipo se ajustan a la especificación más rígida del W3C. Significa que no se pueden usar determinadas etiquetas y atributos HTML que han quedado obsoletos o en desuso en anteriores versiones de HTML. La especificación HTML 4.01 recomienda separar por completo el contenido del documento de su formato. Por lo que, para cumplir escrupulosamente esta recomendación, deberá utilizarse hojas de estilo para especificar el formato y la apariencia de la página. Este tipo de documentos deberán incorporar la siguiente línea al principio del código:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
```



- **Documento HTML transicional.** Los documentos de este tipo se ajustan a la especificación más permisiva del W3C. Significa que se pueden usar determinadas etiquetas y atributos HTML que están en fase de quedarse obsoletas. Esta declaración de tipo de documento es adecuada si se desea combinar la definición del contenido y del formato en el mismo documento HTML. Estos documentos deberán incorporar el siguiente código al principio:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
```

- **Documento HTML con marcos.** Los documentos de este tipo son los que incorporan marcos (etiquetas <frame> y <frameset>). Actualmente, el uso de marcos para componer páginas web está en retroceso, debido a sus limitaciones en cuanto al cumplimiento de las normas de Accesibilidad Web. Estos documentos deberán incorporar el siguiente código:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"
"http://www.w3.org/TR/html4/frameset.dtd">
```

Cuando se emplea la especificación HTML5, la declaración del tipo de documento es más sencilla:

```
<!DOCTYPE html>
```

**Ejemplo.** Crea un documento HTML con Notepad++, denominado *MiPrimerDocumentoHTML.html*, que contenga el siguiente código HTML y comprueba el resultado obtenido en el navegador.

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html lang="es">
  <head>
    <title> Mi primer documento HTML </title>
  </head>
  <body>
    <h1>El lenguaje HTML es fácil de aprender</h1>
    <p> Bienvenido al estudio de HTML. Una buena forma de comprender el
      funcionamiento del lenguaje y conocer con fluidez sus etiquetas es
      practicando su uso mediante la realización de ejercicios. </p>
    <p> Como todos los lenguajes, se aprende mejor cuanto más tiempo se dedique
      a practicar su uso</p>
  </body>
</html>
```

El resultado obtenido en la ventana de visualización del navegador será similar al siguiente:

## El lenguaje HTML es fácil de aprender

Bienvenido al estudio de HTML. Una buena forma de comprender el funcionamiento del lenguaje y conocer con fluidez sus etiquetas es practicando su uso mediante la realización de ejercicios.

Como todos los lenguajes, se aprende mejor cuanto más tiempo se dedique a practicar su uso

## La cabecera

Se utiliza la etiqueta `<head>` para especificar la sección de cabecera del documento HTML. La cabecera de un documento HTML especifica información relativa a la propia página Web y diversos aspectos de configuración del propio documento o del navegador. Entre los diversos aspectos que pueden definirse en la cabecera del documento cabe destacar: el título de la página, la codificación de caracteres que emplea la página y la definición de información adicional de la página.

La **etiqueta `<title>`** permite especificar el título de la página que aparece en la barra superior de la ventana del navegador. Por ejemplo:

```
<title>Ejemplo 1</title>
```

La **etiqueta `<meta>`** permite definir pares identificador-valor que facilitan información o propiedades específicas de un documento HTML, tales como: el autor, la fecha de creación o modificación, la lista de palabras clave relacionadas con el contenido del documento, etc. Esta información adicional, que no se presenta en la ventana de visualización del navegador, puede tener diversas utilidades. La etiqueta `<meta>` no tiene etiqueta de cierre. A continuación, se muestran algunos ejemplos de su uso:

```
<meta name="author" content="Vicente">
<meta name="copyright" content="Vicente">
<meta name="description" content="Curso de HTML">
<meta name="keywords" content="Ciclos, Formativos, Informática, FP, DAW">
<meta name="date" content=" - Fecha de creación: 3 de septiembre de 2016">
```

El valor del identificador *description* suele ser utilizado por los motores de búsqueda para indexar la página, de manera que suele aparecer como resultado de la búsqueda realizada por los usuarios en Internet. La lista de palabras definidas en el valor del identificador *keywords* permiten que la página sea localizable en Internet. La inclusión de palabras no relacionadas con el contenido del documento con el fin de mejorar el posicionamiento Web, es una práctica poco recomendable porque es considerado como spam y puede ser penalizado por los motores de búsqueda.

Se utiliza, también, la etiqueta `<meta>` de cabecera para especificar la **codificación de caracteres** que utiliza la página. Este es un aspecto importante, porque la codificación de caracteres influye de manera directa en que la información contenida en la página se presente de manera adecuada o no en la ventana de visualización del navegador. Si el navegador utilizado está configurado con una codificación de caracteres diferente a la que se ha empleado para construir un documento HTML, entonces los resultados de la presentación sobre el navegador no serán los esperados porque existe una codificación de caracteres inadecuada. En especial, se verá afectada la representación de los caracteres especiales, como, por ejemplo: ñ, á, ò, €, ¿, etc. Para evitar este problema y así, garantizar que la página se presentará siempre de forma adecuada, se recomienda incorporar en el contenido del documento la definición de la codificación de caracteres que utiliza. La codificación de caracteres que utiliza el documento HTML se especifica en la sección de cabecera del documento HTML utilizando el siguiente formato:

```
<meta http-equiv="Content-Type" content="text/html; charset=codificacion">
```

El valor del atributo `charset` debe coincidir con el empleado en el editor de código al construir la página. Las codificaciones de caracteres más usuales, que soportan la totalidad de los caracteres especiales en español, son la codificación **utf-8**, que es compatible a nivel mundial y la codificación **iso-8859-1**, que es compatible a nivel de Europa occidental. Por ejemplo, para definir que un documento HTML está codificado empleando la codificación utf-8 se escribe en la cabecera:

```
<meta http-equiv="Content-Type" content="text/html; charset=utf-8">
```

Cuando se utiliza la especificación HTML5, para definir que un documento HTML está codificado empleando la codificación de caracteres utf-8 se escribe en la cabecera:

```
<meta charset="utf-8" />
```

Además de las etiquetas `<title>` y `<meta>`, en la cabecera pueden aparecer otras etiquetas como pueden ser: `<style>`, `<link>`, etc. que se estudiarán en próximos temas.

### El cuerpo

Se utiliza la etiqueta `<body>` para especificar la sección de cuerpo del documento HTML. En el cuerpo del documento HTML se especifican los elementos que incorporan la información que se presentará en la ventana de visualización del navegador. Por lo tanto, el cuerpo define el contenido del documento HTML. La mayoría de las etiquetas de HTML especifican elementos que pueden aparecer en el cuerpo, solo unas pocas etiquetas pueden aparecer en la cabecera. En los próximos apartados y temas se irán estudiando las etiquetas de los diversos elementos que pueden aparecer en el cuerpo del documento.

Los principales atributos de la etiqueta `<body>` son los siguientes:

- **id.** Sirve para asignar un nombre único al elemento. Se utiliza para identificar elementos en hojas de estilo y en scripts.
- **class.** Sirve para asignar un nombre de clase de estilo. Es posible usar el mismo nombre de clase para varios elementos. Se usa para mejorar el rendimiento de las hojas de estilo.
- **title.** Sirve para agregar un comentario asociado a un elemento. Los navegadores muestran este comentario en una ventana emergente cuando el ratón se sitúa sobre el elemento.
- **style.** Sirve para aplicar características de estilo al elemento.

Existen atributos del elemento `<body>` que han sido muy utilizados en el pasado, pero que las últimas versiones de HTML se han declarado en desuso y, por tanto, no debería utilizarse. Estos atributos en desuso son: `background`, `text`, `link`, `alink`, `vlink`, `bgcolor`.

En las últimas versiones de HTML, la forma válida de añadir color al fondo del documento se realiza empleando las características de las hojas de estilo. Aunque las hojas de estilo se estudiarán en temas posteriores con profundidad, en el siguiente ejemplo se adelanta cómo cambiar el color del fondo para mejorar la apariencia de la página. En el ejemplo, puede observarse el uso del atributo

`style` en la etiqueta `<body>` para definir el color de fondo, utilizando para ello las características de las hojas de estilo. También puede comprobarse que se ha reescrito el código del ejemplo anterior, denominado *MiPrimerDocumentoHTML.html*, utilizando una versión más moderna de HTML.

```
<!DOCTYPE html>
<html lang="es">
<head>
  <title>Mi primer documento HTML - Mejorado</title>
  <meta charset="utf-8" />
</head>
<body style="background-color:LightSkyBlue">
  <h1>El lenguaje HTML es fácil de aprender</h1>
  <p>
    Bienvenido al estudio de HTML. Una buena forma de comprender el
    funcionamiento del lenguaje y conocer con fluidez sus etiquetas es
    practicando su uso mediante la realización de ejercicios.
  </p>
  <p>
    Como todos los lenguajes, se aprende mejor cuanto más tiempo se dedique
    a practicar su uso
  </p>
</body>
</html>
```

## 2.5 ELEMENTOS BÁSICOS

En este apartado se estudian los elementos más básicos que pueden encontrarse dentro del cuerpo de un documento HTML. En el lenguaje HTML se pueden diferenciar dos tipos de elementos de la página: los elementos de bloque y los de línea.

- **Elementos de bloque.** Son elementos que crean estructuras más grandes que pueden contener elementos de línea y otros elementos de bloque. La característica principal de este tipo de elementos es que siempre incorporan un salto de línea antes y después del elemento. Ejemplos de elementos de bloque son: `<p>`, `<div>`, `<table>`, `<ul>`, etc.
- **Elementos de línea.** Estos elementos pueden contener información a presentar y, también pueden contener a otros elementos de línea. Estos elementos se reconstruyen en la ventana de visualización del navegador a continuación del elemento anterior, sin introducir ningún salto de línea. Ejemplos de elementos de línea son: `<span>`, `<small>`, `<b>`, etc.

El texto es el tipo de contenido más habitual en las páginas web. La correcta estructuración del texto facilitará la lectura de la página por parte de otras personas. Cuando se escribe texto en un documento HTML es necesario tener presente que los espacios en blanco consecutivos entre palabras se presentan en la ventana de visualización del navegador como un único espacio en blanco. Y que, además, los retornos de carro (Tecla `<Return>`) introducidos en el código no aparecen representados visualmente cuando se presenta la página. Esto es debido a que el lenguaje HTML se creó para presentar texto en cualquier situación, por lo que siempre debe utilizarse la etiqueta correspondiente para representar cualquier efecto en la presentación visual del texto.

A continuación, se estudian las etiquetas del lenguaje HTML más importantes para organizar el texto en una página.

### Párrafo

El párrafo es la estructura más utilizada. La etiqueta de definición de un elemento párrafo es `<p>`. Un elemento párrafo se compone de una etiqueta de apertura, el contenido a presentar y una etiqueta de cierre. Para definir que un texto pertenece al mismo párrafo, se introduce dentro de las etiquetas `<p>` `</p>`. Cuando el párrafo termina con la etiqueta de cierre, se introduce automáticamente un salto de línea. Por ejemplo:

```
<!DOCTYPE html>
<html lang="es">
<head>
  <title>Párrafo</title>
  <meta charset="utf-8" />
</head>
<body>
  <p>Este texto forma parte del párrafo 1.</p> <p>Este texto forma parte del párrafo 2.
  Por eso está en otra línea.</p>
</body>
</html>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte del párrafo 1.

Este texto forma parte del párrafo 2. Por eso está en otra línea.

El elemento párrafo es un elemento de bloque y no puede contener otros elementos de bloque. El navegador se encarga de que la línea de texto ocupe todo el ancho disponible, y si el ancho cambia, el texto se ajusta al nuevo ancho ocupando más o menos líneas. En un documento HTML, cualquier texto debe formar parte de un elemento, es decir, no puede estar fuera de una etiqueta.

### Salto de línea

La etiqueta `<br>` fuerza un salto de línea. Esta etiqueta no tiene etiqueta de cierre. Por ejemplo:

```
<!DOCTYPE html>
<html lang="es">
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
  <title>Salto de línea</title>
  <meta charset="utf-8" />
</head>
<body>
  <p>
    Este
    texto
    aparece en
    la misma línea.
  </p>
```

```
Este <br> texto<br>
aparece <br> en <br>
líneas <br> distintas.
</p>
</body>
</html>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto aparece en la misma línea. Este  
texto  
aparece  
en  
líneas  
distintas.

Esta etiqueta no debe usarse más que cuando sea necesario. En su lugar se suele agrupar el texto en párrafos, tal como veremos en el siguiente punto.

## Encabezados

Los encabezados son bloques que destacan el texto que contienen. Se utilizan como delimitadores de sección dentro de las páginas web. Existen 6 grados o niveles de encabezados, numerados del 1 al 6, siendo el 1 el más grande y el 6 el más pequeño. Su sintaxis es la siguiente:

`<h> Texto del encabezado </h>`

Siendo n el número de grado o nivel de encabezado, entre 1 y 6. Los encabezados, o también denominados títulos, son elementos de bloque. Por ejemplo:

```
<!DOCTYPE html>
<html lang="es">
<head>
  <meta http-equiv="Content-Type" content="text/html; charset=utf-8"/>
  <title>Encabezados</title>
  <meta charset="utf-8" />
</head>
<body>
  <body>
    <h1>Encabezado de nivel 1.</h1>
    <h2>Encabezado de nivel 2.</h2>
    <h3>Encabezado de nivel 3.</h3>
    <h4>Encabezado de nivel 4.</h4>
    <h5>Encabezado de nivel 5.</h5>
    <h6>Encabezado de nivel 6.</h6>
  </body>
</body>
</html>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

**Encabezado de nivel 1.**

**Encabezado de nivel 2.**

**Encabezado de nivel 3.**

**Encabezado de nivel 4.**

**Encabezado de nivel 5.**

**Encabezado de nivel 6.**

Como los encabezados van a definir el esquema del documento, deben ser utilizados de forma adecuada y coherente al propósito del texto. Los encabezados agregan un salto de línea al final. El tamaño de la letra y el tipo de destacado dependen del navegador que se utilice. La etiqueta de encabezado no debe utilizarse dentro de una etiqueta de párrafo y viceversa.

### Línea horizontal

La etiqueta `<hr>` presenta una línea horizontal en la ventana de visualización. Se suele utilizar de separador. La línea horizontal, o también denominada regla horizontal, introduce un salto de línea al final. Por ejemplo:

```
<p>
Este es el primer párrafo de un texto.
</p>
<hr>
<p>
Este es el segundo párrafo que se presenta separado del anterior mediante una línea
que ocupa todo el ancho de la ventana de visualización del navegador.
</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este es el primer párrafo de un texto.

Este es el segundo párrafo que se presenta separado del anterior mediante una línea  
que ocupa todo el ancho de la ventana de visualización del navegador.



### Texto preformateado o con formato previo

La etiqueta <pre> sirve para que el navegador presente un texto tal y como aparece en el contenido de esta etiqueta en el código HTML, es decir, respetando tabuladores, espacios y saltos de línea. Esta etiqueta hace, además, que se muestre el texto con una fuente de ancho fijo. Es útil para presentar texto con una determinada estructura específica. Por ejemplo:

```
<!DOCTYPE html>
<html lang="es">
<head>
  <meta charset="utf-8" />
  <title>Texto preformateado</title>
</head>
<body>
  <h1>Función: mayorEntero()</h1>
  <p>El código de la función que devuelve el mayor de dos números enteros es la siguiente:</p>
  <pre>
    function mayorEntero(int a, int b)
    {
      int m = a;
      if (b > m)
      {
        m = b;
      }
      return (m)
    }
  </pre>
</body>
</html>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

## Función: mayorEntero()

El código de la función que devuelve el mayor de dos números enteros es la siguiente:

```
function mayorEntero(int a, int b)
{
  int m = a;
  if (b > m)
  {
    m = b;
  }
  return (m)
}
```

## Comentarios

Los comentarios se introducen entre los símbolos `<!--` y `-->`. Todo el texto que se coloque entre estos símbolos es ignorado por el navegador, es decir, no será interpretado y, por tanto, no se presentará de ningún modo en la ventana de visualización del navegador. Los comentarios no se pueden anidar. Por ejemplo:

```
<!-- Este texto es un comentario, aunque ocupe varias líneas. No aparecerá en la ventana de visualización del navegador ni afecta a la página web. Se suelen utilizar para realizar anotaciones útiles para el desarrollador, como fechas, explicaciones, etc. -->
```

## Secuencias de escape

Las secuencias de escape son sucesiones de caracteres que representan un carácter determinado de forma universal. Pueden emplearse para presentar en la ventana de visualización del navegador los caracteres reservados que tienen un uso específico para el lenguaje HTML. Así, por ejemplo, los caracteres `'<'`, `'>'`, `'&'`, etc. tienen un significado específico para HTML y, por tanto, no pueden emplearse directamente para ser presentados porque el navegador se podría confundir. Estos caracteres reciben el nombre de caracteres especiales. Si se introduce el carácter `<` en alguna parte de la información a presentar dentro de un elemento de HTML, el navegador interpretará que lo que viene a continuación de ese carácter es una etiqueta.

Para solucionar el problema planteado se utilizan las secuencias de escape que se caracterizan por comenzar por el carácter `'&'` y finalizar con el carácter `';'`. Para introducir los caracteres que tienen un uso específico para HTML en el texto a presentar en la página, pueden emplearse las siguientes secuencias de escape reservados reciben el nombre de caracteres especiales:

Símbolo	Secuencia de escape
©	&copy;
®	&reg;
<	&lt;
>	&gt;
&	&amp;
Espacio	&nbsp;
"	&quot;

Tal como se estudió anteriormente, si se introducen varios espacios en blanco entre dos palabras, se presentarán como un único espacio en blanco en la ventana de visualización del navegador. Si se desea introducir más de un espacio en blanco entre palabras, o bien, si se desea introducir un espacio en blanco en cualquier punto de la página, puede emplearse una secuencia de escape `&nbsp;` (*no-break space*) por cada espacio que se desee añadir.

Por otra parte, las secuencias de escape se pueden emplear para representar los caracteres no ingleses, que son específicos de cada lengua, cuando se emplean juegos de codificación de caracteres de 7 bits. En español, las letras con acento mayúsculas y minúsculas, la eñe mayúsculas y minúsculas, etc. pueden representarse de mediante las siguientes secuencias de escape:

Símbolo	Secuencia de escape
ñ	&ntilde;
Ñ	&Ntilde;
á	&aacute;
Á	&Aacute;
é	&eacute;
í	&iacute;
ó	&oacute;
ú	&uacute;
à	&agrave;
è	&egrave;
È	&Egrave;
ì	&igrave;
ò	&ograve;
ù	&ugrave;
ö	&ouml;
Ö	&Ouml;

Actualmente, no suelen emplearse secuencias de escape para representar caracteres no ingleses, porque se emplean juegos de caracteres basados en codificaciones de 8 bits. De manera que las codificaciones de caracteres UTF-8 o ISO-8859-1 no presentan problemas para representar caracteres no ingleses.

## 2.6 FORMATO DEL TEXTO

En este apartado se estudian algunas etiquetas que sirven para dar forma al texto mediante elementos de control de la tipografía: cursiva, negrita, subrayado, tipos y tamaños de letra, etc.

Es muy importante tener en cuenta que, actualmente, la mayoría de las etiquetas de HTML que se emplean para dar formato al texto han quedado obsoletas en las especificaciones más modernas de HTML. En efecto, el planteamiento actual que se sigue en las últimas versiones del lenguaje consiste en utilizar HTML solo para organizar el texto, y no para dar formato visual ni mejorar la estética de la página. De modo que, la apariencia de la página se especifica, actualmente, utilizando las hojas de estilo en cascada que se estudiarán en temas posteriores. Las etiquetas obsoletas se incluyen para facilitar su identificación y sólo se utilizarán en los ejercicios de esta sección hasta que se estudien las hojas de estilo en cascada.

Todas las etiquetas que se refieren al formato del texto son elementos de línea, por lo que no incorporan salto de línea y se pueden introducir dentro de otros elementos de bloque o de línea.

Las etiquetas de formato de texto pueden combinarse para producir varios efectos a la vez. Por ejemplo, se puede subrayar y poner en negrita un texto o poner en cursiva y cambiar el tipo de letra, etc. En estos casos, solo es necesario cerrar las etiquetas de manera correcta, es decir, inversamente a como se han abierto, tal como se indicó anteriormente.

### Negrita

Para resaltar una parte del texto en negrita se utiliza la etiqueta `<b>`. Esta etiqueta se aplica al texto que está especificado dentro de otro elemento, como puede ser un párrafo. Por ejemplo:

```
<p>Este texto forma parte de un párrafo. Y esta <b>palabra</b> está en <b>negrita</b> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta **palabra** está en **negrita** dentro del mismo párrafo.

Esta etiqueta tiene un sinónimo. En lugar de utilizar la etiqueta `<b>` se puede también utilizar también la etiqueta `<strong>`, obteniéndose idéntico resultado.

```
<p>Este texto forma parte de un párrafo. Y esta <strong>palabra</strong> está en <strong>negrita</strong> dentro del mismo párrafo.</p>
```

### Cursiva

Para resaltar una parte del texto en cursiva, se utiliza la etiqueta `<i>`. Por ejemplo:

```
<p>Este texto forma parte de un párrafo. Y esta <i>palabra</i> está en <i>cursiva</i> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta *palabra* está en *cursiva* dentro del mismo párrafo.

Como puede apreciarse en el ejemplo anterior, esta etiqueta está incluida en un elemento párrafo, de modo que se muestra en cursiva una parte del texto del párrafo.

Esta etiqueta tiene un sinónimo. En lugar de utilizar la etiqueta `<i>` se puede también utilizar también la etiqueta `<em>`, obteniéndose idéntico resultado.

### Texto monoespaciado

Para que una parte del texto se presente con letras monoespaciadas, también denominadas de tipo teletipo se utiliza la etiqueta <tt>. De este modo, se presenta el texto con un ancho fijo, es decir, cada letra ocupa el mismo espacio. Por ejemplo:

```
<p>Este texto forma parte de un párrafo. Y esta <tt>palabra</tt> está en <tt>modo teletipo</tt> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta palabra está en modo teletipo dentro del mismo párrafo.

### Tamaño de letra grande

Para resaltar una parte del texto con un tamaño de letra mayor se utiliza la etiqueta <big>. Por ejemplo:

```
<p>Este texto forma parte de un párrafo. Y esta <big>palabra</big> está en <big>tamaño grande</big> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta palabra está en tamaño grande dentro del mismo párrafo.

### Tamaño de letra pequeño

Para resaltar una parte del texto con un tamaño de letra más pequeño se utilizan las etiquetas <small> </small>. Estas etiquetas pueden (y deben) utilizarse dentro de las etiquetas de párrafo. Por ejemplo, el código siguiente:

```
<p>Este texto forma parte de un párrafo. Y esta <small>palabra</small> está en <small>tamaño pequeño</small> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta palabra está en tamaño pequeño dentro del mismo párrafo.

A continuación, se estudian las etiquetas obsoletas y que, por tanto, no deberían ser utilizadas para crear nuevos documentos HTML: <u>, <s> y <font>.

## Subrayado

Para subrayar una parte del texto se utiliza la etiqueta <u>. En el sistema de información Web no suele emplearse el subrayado para resaltar el texto, se suele preferir utilizar texto en negrita o en cursiva. Ello es debido a que los enlaces o hipertextos que se presentan en las páginas suelen presentarse como texto subrayado, de manera que los usuarios suelen pensar que cuando se presenta un texto subrayado es porque se trata de un enlace. Por ejemplo:

```
<p>Este texto forma parte de un párrafo. Y esta <u>palabra</u> está <u>subrayada</u> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta palabra está subrayada dentro del mismo párrafo.

## Tachado

Para tachar una parte del texto con una línea a media altura se utiliza la etiqueta <s> o <strike>. Por ejemplo, el código siguiente:

```
<p>Este texto forma parte de un párrafo. Y esta <s>palabra</s> está <strike>tachada</strike> dentro del mismo párrafo.</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Este texto forma parte de un párrafo. Y esta ~~palabra~~ está ~~tachada~~ dentro del mismo párrafo.

## Cambiar el tipo, color y tamaño de letra

Para cambiar el tipo, color o tamaño de la fuente de letra de una parte del texto se utiliza la etiqueta <font>. Para ello, se utilizan los siguientes atributos:

- size. Permite cambiar el tamaño de la letra. Puede tomar valores enteros entre 1 y 7. El tamaño de letra establecida por defecto corresponde con el valor numérico 3.
- face. Permite cambiar el tipo o la familia de la fuente. Los valores que toma se corresponden con un tipo de letra: "Arial", "Tahoma", "Helvetica", etc.
- color. Permite cambiar el color del texto. Puede tomar valores hexadecimales precedidos por el signo # seguido de tres parejas de dígitos hexadecimales que siguen la especificación RGB. También admite nombres predefinidos de colores: "red", "black", "yellow", etc.

Se puede utilizar uno, dos o tres atributos a la vez dentro de la misma etiqueta <font>. Debe tenerse en cuenta que los valores de los atributos se escriben siempre entre comillas. Por ejemplo:

```
<p>  
<font color="green" face="Arial">Esta frase está en color verde y letra Arial.</font>  
<font color="#FF0000" size="5">Esta frase está en color rojo y tiene tamaño 5.</font>  
<font face="Verdana">Y esta frase utiliza la fuente Verdana.</font>  
</p>
```

El ejemplo anterior, produce el siguiente resultado en la ventana de visualización del navegador:

Esta frase está en color verde y letra Arial. Esta frase está en color rojo y tiene tamaño 5. Y esta frase utiliza la fuente Verdana.

## 2.7 VALIDACIÓN DEL CÓDIGO HTML

Para verificar que el código de un documento HTML o una página Web es correcto, se puede consultar la dirección <http://validator.w3.org/>. Esta dirección Web accede a un validador de código HTML, en el que se puede insertar el código del documento HTML que se desee comprobar o, bien, se puede subir el archivo que contiene el código HTML a comprobar. Tras realizar la comprobación, ofrecerá un resultado sobre los errores de validación que contiene.

Para comprobar el código de una página web, dicho código deberá contener obligatoriamente una declaración de tipo de documento y tener definida la codificación de caracteres que utiliza.



## Glosario de términos

**Hojas de estilo en cascada o CSS (*Cascading Style Sheets*).** Es un lenguaje usado para definir la apariencia de la presentación de un documento estructurado escrito en HTML o XML. El *World Wide Web Consortium* (W3C) es el organismo encargado de formular la especificación de las hojas de estilo que sirven de estándar para los navegadores. La idea que se encuentra detrás del desarrollo de las hojas de estilo es separar la estructura de un documento de su presentación.

**HTML (*HyperText Markup Language*).** Lenguaje de marcado predominante para la elaboración de páginas Web que se utiliza para especificar la presentación y la información contenida. El código HTML se escribe en forma de etiquetas o marcas, rodeadas por corchetes angulares (<, >).

**Interfaz de usuario (*User Interface*).** Procesamiento (programa informático) que actúa empleando un conjunto de imágenes y objetos para presentar la información y las acciones disponibles que puede realizar el usuario. Su principal uso, consiste en proporcionar un entorno visual que facilite la comunicación del usuario con una aplicación informática o con el sistema operativo, de modo que transformen un sistema automatizado no humano en una herramienta convivencial. En la actualidad, dado que la mayor parte de las interfaces de usuario son de tipo gráfico, éstas suelen recibir la denominación de *Graphical User Interface* (GUI) o interfaz gráfica de usuario.

**URL (*Uniform Resource Locator*).** Es un identificador de recursos uniforme (*Uniform Resource Identifier, URI*) cuyos recursos a los que se refiere pueden cambiar, es decir, la dirección puede apuntar a recursos variables en el tiempo. Están formados por una secuencia de caracteres, de acuerdo a un formato estándar que designa recursos en una red, como puede ser Internet. Las URL fueron usadas por primera vez por Tim Berners-Lee en 1991, para permitir a los autores de documentos establecer hipervínculos o enlaces en la *World Wide Web* (WWW). Desde 1994, en los estándares de Internet, el concepto de URL ha sido incorporado dentro del más general de URI, pero el término URL todavía se utiliza ampliamente.

**XHTML (*eXtensible HyperText Markup Language*).** Es, básicamente, el lenguaje HTML aunque expresado como XML válido. En su versión 1.0, XHTML es únicamente la versión XML de HTML, por lo que tiene, básicamente, las mismas funcionalidades, pero cumple unas especificaciones más estrictas que HTML. Su objetivo es avanzar en el proyecto del *World Wide Web Consortium* (W3C) de lograr una web semántica, donde la información, y la forma de presentarla estén claramente separadas. La versión 1.1 de XHTML es similar, pero partiendo a la especificación en módulos.

**World Wide Web (WWW).** También se denomina Web de forma simplificada. La *World Wide Web*, o red informática mundial, es un sistema de distribución de documentos de hipertexto e hipermedia interconectados y accesibles vía Internet. Utilizando un navegador web, un usuario puede visitar sitios web compuestos de páginas web que pueden contener texto, imágenes, vídeos u otros contenidos multimedia, y puede navegar a través de esas páginas web usando hiperenlaces.