

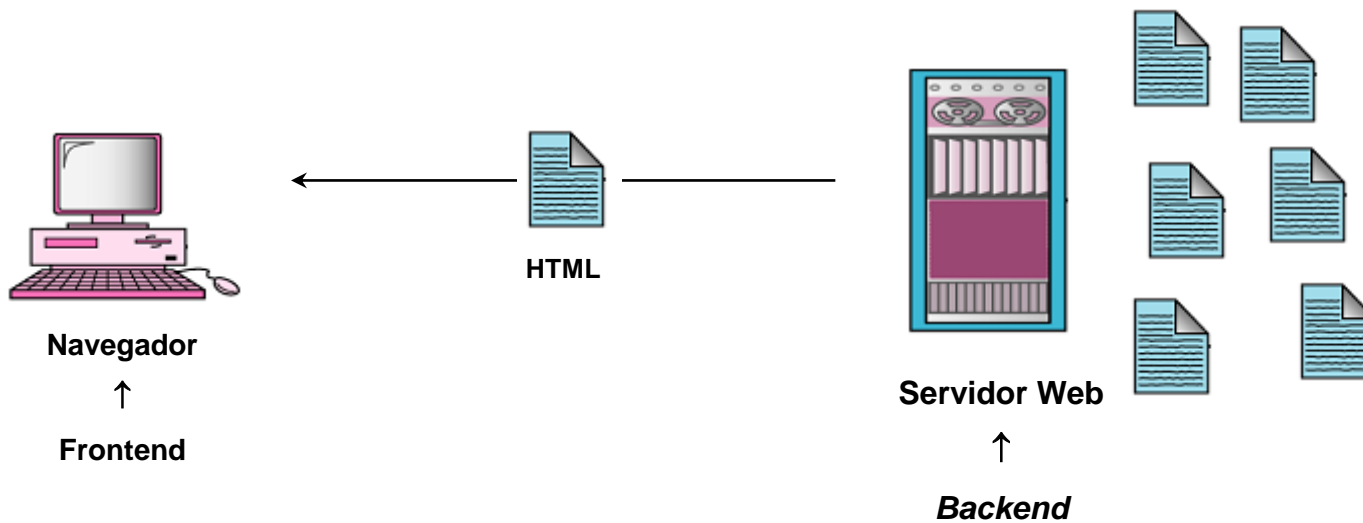


# Desarrollo de Aplicaciones Web

*HTML. Introducción.*

# La WEB

Conjunto de documentos residentes en diferentes máquinas, relacionados mediante hiperenlaces.



*Frontend* de un sitio web → Todo aquello con lo que el usuario interactúa.

# La WEB: *Frontend*

**La Página** → Documento que incluye el contenido a mostrar

**HTML** (HyperText Markup Lenguaje). Propuesto en 1989:

- Universalidad
- Lenguaje de marcas → etiquetas (*Tunnicliffe, 1967*)
- Hipertexto (*Ted Nelson, 1965*)

```
<html>
  <head>
    <title>Desarrollo de Aplicaciones Web</title>
  </head>
  <body>
    <h1>Bienvenidos a la página web de DAW</h1>
    <p>Esta es la página web de la asignatura ...
    </p>

    <p>Aquí encontraras todo el material ...
    </p>
  </body>
</html>
```

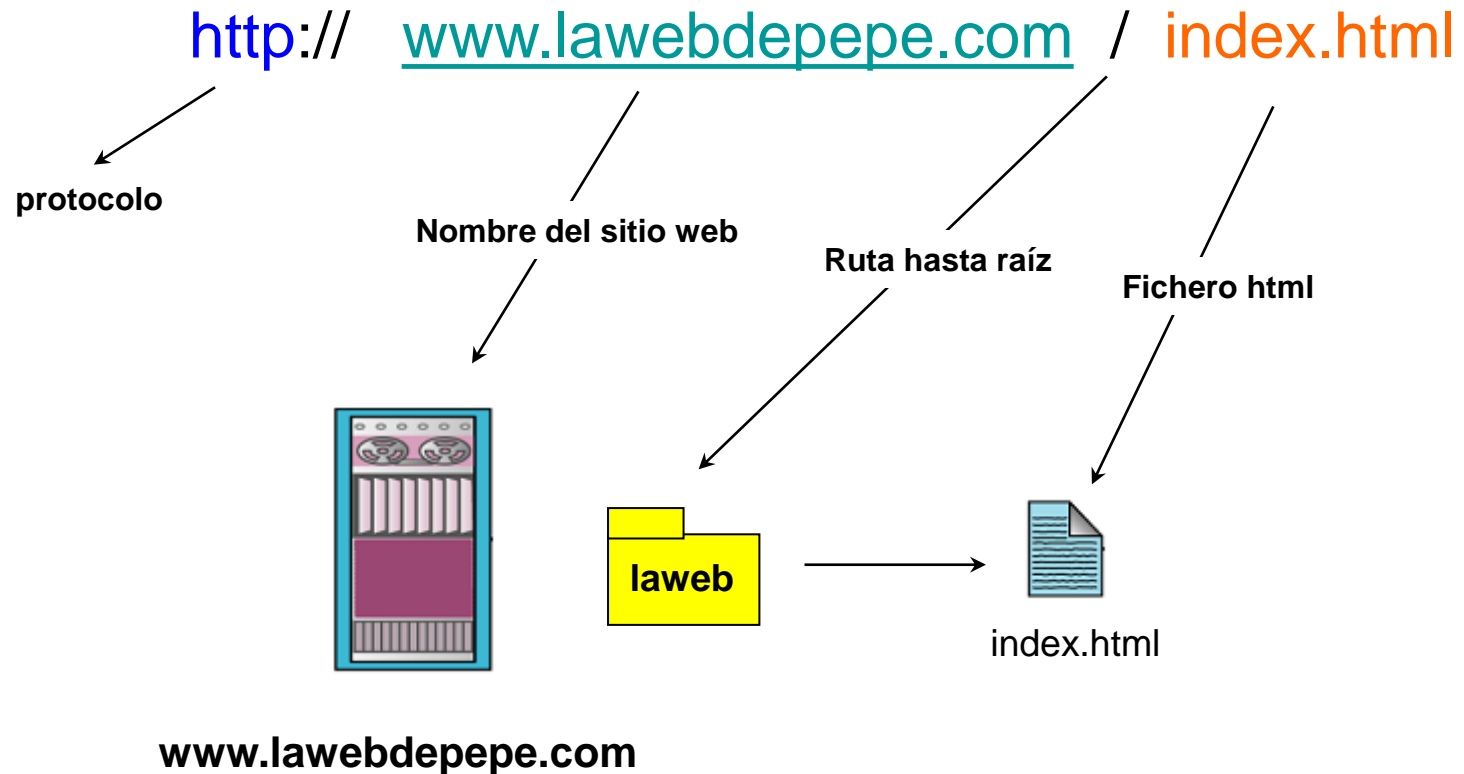
# La WEB: *Frontend*

**El navegador** → Herramienta que interpreta el código html escrito en la página web. Actuaría como un cliente.

**Motor de renderizado** de un navegador web → parte del navegador encargada de tomar el contenido (HTML, XML, imágenes, etc.) y la información de formato (CSS, etc.) y crear una representación visual de la página.

- **Antes 1993.**- navegadores incluían poca funcionalidad (*WorldWideWeb* → *Nexus*)
- **1993 Marc Andreessen.**- desarrolla el *Mosaic* en la NCSA
- **1994 surge Netscape.**- incluye extensiones HTML propias que permiten ver texto coloreado, fotografías, etc. Netscape domina el mercado. Microsoft contra-ataca con Internet Explorer. Guerra de los navegadores
- **1994.**- se crea el consorcio [W3C](#) (Tim Berners-Lee), con el fin de poner orden y establecer estándares

# Publicación de una página web:



# Publicación de una página web:

- Insercción en servidor web propio
- Hosting

## Nombre de dominio:

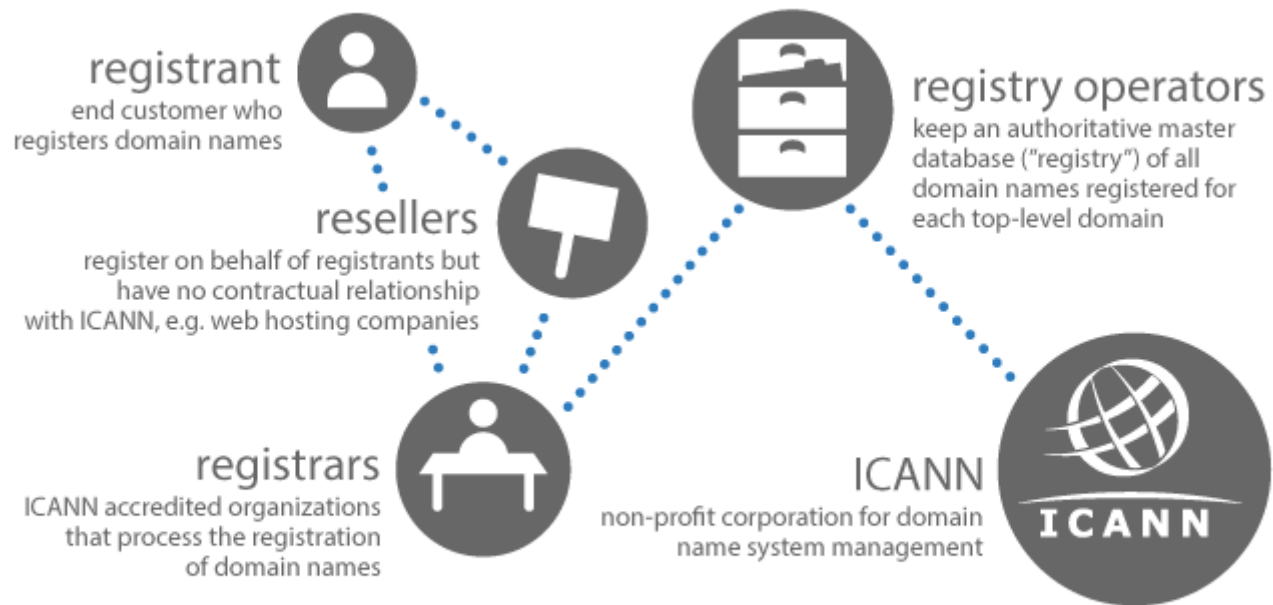
<http://www.usc.es/lawebdepepe> → <http://www.lawebdepepe.com>

**ICANN** → Internet Corporation for Assigned Names and Numbers.

**WHOIS** → Directorio que contiene toda la información técnica y de contacto de los propietarios de los nombres de dominio.

**DNRD** → Domain Name Registration Data

# Publicación de una página web:



domain registry process

[Registrars Accredited](#)

# Historia del HTML

- **1989** .- *Tim Berners-Lee* y *Robert Cailliau* presentan su propuesta de hipertexto a una convocatoria del [CERN](#) que ganan. La propuesta la denominaron [WorldWideWeb \(W3\)](#).
- **1991**.- Se publica el primer documento formal con la descripción del *HTML*: “[HTML Tags](#)”, que incluía 22 etiquetas.
- 1993**.- Primera propuesta oficial para convertir *HTML* en un estándar por parte de [Internet Engineering Task Force \(IETF\)](#). No llega a fructificar.
- 1996**.- El *World Wide Consortium (W3C)* publicará a partir de aquí los diferentes estándares de *HTML* que van apareciendo.
- **2004**.- *Apple*, *Mozilla* y *Opera* crean el [Web Hypertext Application Technology Working Group \(WHATWG\)](#), para relanzar la estandarización del *HTML* ante la pasividad de *W3C*.
- **Enero 2008**.- Se publica el primer borrador (*Working Draft*) del *HTML 5.0*.
- **Diciembre 2012**.- Se publica la primera recomendación de candidato (*Candidate Recommendation*) del [HTML 5.0](#).
- **Julio 2014**.- *W3C* comienza a discutir la introducción de los *Web Components*. Permiten la creación de etiquetas personalizadas por parte del desarrollador.



# HTML. Proceso de estandarización

## - HTML 1.0, 2.0

- No hay lugar para los adornos. Texto plano con hiperenlaces

## - HTML 3.0

- Aparecen nuevas etiquetas propietarias.
- Posibilidad de contenidos enriquecidos mediante colores e imágenes.
- Guerra de navegadores.

## - HTML 3.2:

- Se estandarizan extensiones propietarias, incluyendo algunas y eliminando otras.
- Se anima a los creadores de navegadores a seguir dichos estándares.

## - HTML 4.0:

- Se separa contenido de estructuras y formato (CSS).

## - HTML 4.01:

- Todo elemento de línea, tiene que estar dentro de un elemento de bloque

## - XHTML 1.0:

- Se busca mantener la estructura XML.
- Las etiquetas que abren, necesariamente han de cerrarse.

## - HTML 5.0:

- Menor compatibilidad con versiones obsoletas → Facilita desarrollo de nuevos navegadores
- Nuevas etiquetas.

# HTML. Versiones

## **Estricta:**

- Solo se permite el uso de las etiquetas actualmente aprobadas.

## **Transitoria:**

- Permite el uso de etiquetas no aprobadas.

## **Conjunto de Marcos:**

- Permite el uso de etiquetas no aprobadas.
- Permite el uso de marcos (desaconsejados por W3C en la actualidad).

# HTML. Versiones.

**<!DOCTYPE .....>**

**<html>**

**<head>**

**<title>Programación Básica en Internet</title>**

**</head>**

**<body>**

**<h1>Bienvenidos a la página web de PBI</h1>**

**<p>Esta es la página web de la asignatura ...**

**</p>**

**</body>**

**</html>**

# HTML. Versiones

## HTML 4.01

### - Estricta:

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

### - Transitoria:

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

### - Conjunto de Marcos:

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

## XHTML 1.0

### -Estricta:

-<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">

### - Transitoria:

- <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">

### - Conjunto de Marcos:

- <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">

W3C proporciona mecanismos para validar páginas

<http://validator.w3.org>

# HTML. Codificación.

HTML pretende ser un documento en formato texto universal. Es necesario declarar el conjunto de símbolos utilizados → Uso de la etiqueta *<meta>* dentro de head.

```
<html>
  <head>
    <meta http-equiv="content_type" content="text/html; charset= [encoding]">
    <title>Desarrollo de Aplicaciones Web</title>
  </head>

  <body>
    <h1>Bienvenidos a la página web de DAW</h1>
    <p>Esta es la página web de la asignatura ...
    </p>
  </body>
</html>
```

# HTML. Codificación

Los valores para *encoding* aparecen en la lista de W3C:

<http://www.w3.org/International/O-Charset-lang.html>

ejemplo: [iso-8859-1](#) para español.

Dichos valores se encuentran obsoletos en la actualidad. Se recomienda el uso de codificación *utf-8*

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

# HTML. Conseguir que nos visiten

Existen más de 2 billones de páginas web y se siguen creando miles cada día.  
Cómo coseguir que nos **visiten**???

- Uso de palabras clave:
  - En el título: `<title> .....</title>`
  - En los encabezados: `<h1> ..... </h1>`
  - En las etiquetas `<alt>` de imagen
- Uso de palabras clave explícitas:
  - Etiquetas meta de head: `<meta name="keywords" content= "..., ..., ...">`
- Inclusión de descripción explícita:
  - Etiquetas meta de head: `<meta name="description" content= ".....">`
- Envío del sitio web a un motor de búsqueda:
  - [Google](#)

# HTML. Directivas para Robots de Búsqueda.

Directiva	Significado
<b>all</b>	No hay restricciones de indexación ni de presentación de contenido. Nota: Esta directiva es el valor predeterminado y no tiene ningún efecto si se muestra de forma explícita.
<b>noindex</b>	No se muestra ni esta página ni un enlace "en caché" en los resultados de búsqueda.
<b>nofollow</b>	No se siguen los enlaces de esta página.
<b>none</b>	Equivalente a noindex, nofollow.
<b>noarchive</b>	No se muestra ningún enlace "en caché" en los resultados de búsqueda.
<b>nosnippet</b>	No se muestra ningún fragmento en los resultados de búsqueda de esta página.
<b>noodp</b>	No se utilizan metadatos del <a href="#">proyecto de Open Directory</a> para los títulos o los fragmentos que se muestran en esta página.
<b>notranslate</b>	No se ofrece una traducción de esta página en los resultados de búsqueda.
<b>noimageindex</b>	No se indexan las imágenes de esta página.
<b>unavailable_after: [RFC-850 date/time]</b>	No se muestra esta página en los resultados de búsqueda después de la fecha y la hora especificadas. La fecha y la hora deben especificarse en el <a href="#">formato RFC 850</a> .



# HTML. Directivas para Robots de Búsqueda.

`<meta name="robots" content="noindex, nofollow" />`

## Robots de Búsqueda

- Googlebot → Google
- MSNBot → Microsoft

Terms	Googlebot	MSNBot
NoIndex	<u>YES</u>	<u>YES</u>
NoFollow	<u>YES</u>	<u>YES</u>
NoArchive	<u>YES</u>	<u>YES</u>
NoSnippet	<u>YES</u>	NO
NoODP	<u>YES</u>	<u>YES</u>
NoYDIR	NO	NO
NoImageIndex	<u>YES</u>	NO
NoTranslate	<u>YES</u>	NO
Unavailable_After	<u>YES</u>	NO

# HTML. Herramientas de Análisis.

Navegador	Herramienta
Apple Safari	Safari's Web Inspector
Google Chrome	Google's Chrome Developers Tools (CTRL+Shift+J)
Microsoft IE	Internet Explorer's Developer Tools
Mozilla Firefox	Firebug, Firefox Web Developer
Opera	Dragonfly



# Desarrollo de Aplicaciones Web

*HTML. Introducción.*