

Introducción a los estándares de la Web

Lenguajes para Internet
Curso 2007/2008

Juan Manuel Cueva Lovelle
cueva@lsi.uniovi.es
www.di.uniovi.es/~cueva

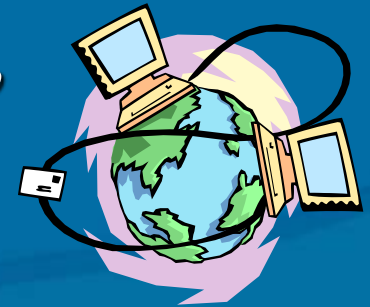
Departamento de Informática
Universidad de Oviedo (Asturias, España)

OOTLab www.ootlab.uniovi.es



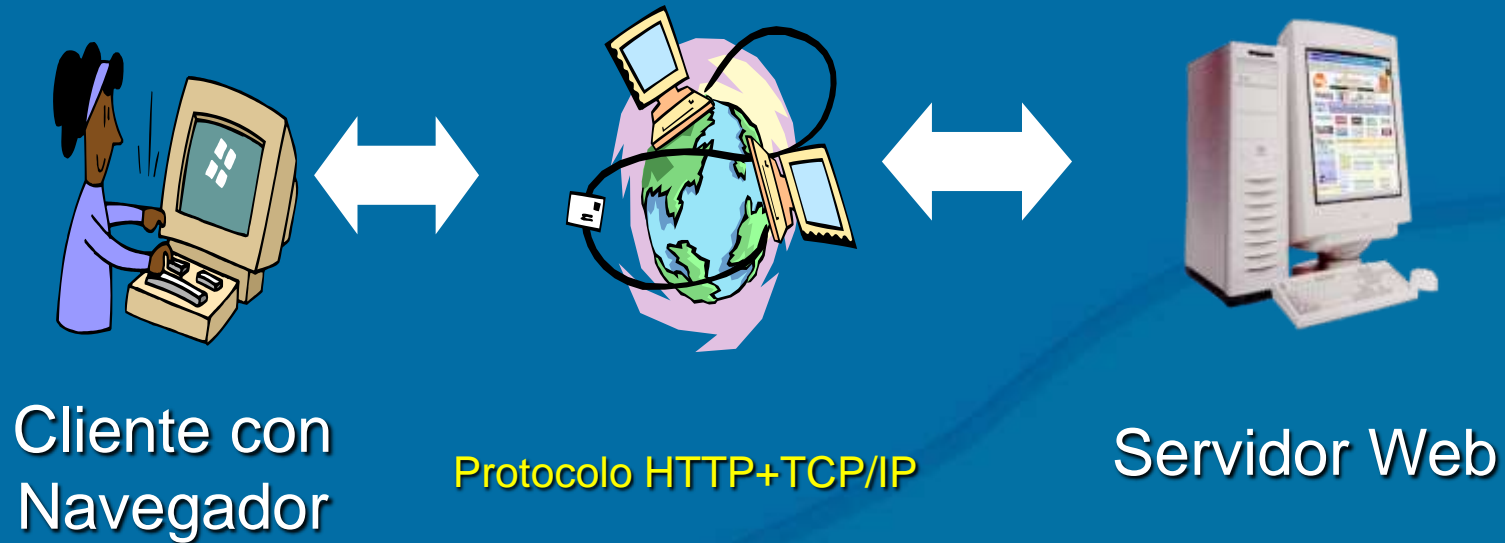
¿Qué es la Ingeniería Web?

- ¿Qué es la Web?
 - Literalmente "telaraña"
 - Es la denominación de la red Internet *World-Wide Web*
 - Red de Ordenadores conectados mediante protocolo TCP/IP + HTTP (Tim Berners-Lee, 1990)
- Ingeniería Web
 - *"Es el proceso para crear, implantar y mantener aplicaciones y sistemas Web de alta calidad"*



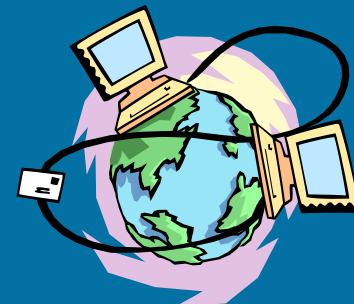
¿Cómo funciona la Web?

Arquitectura cliente-servidor



Estructura básica de Internet

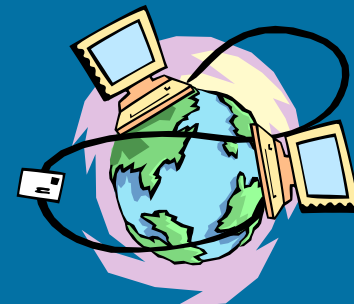
Protocolo TCP/IP



- **Protocolo:** Es un conjunto de reglas que arbitra el intercambio de información entre dos computadoras y se debe seguir estrictamente para la comunicación entre ambos.
- El protocolo **TCP/IP** (*Transmission Control Protocol / Internet Protocol*) es una arquitectura de varios niveles en la que los protocolos de cada nivel tienen un cometido específico sobre el que se apoyan los protocolos de niveles superiores

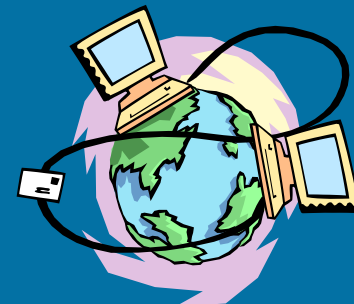
Niveles TCP/IP

Esquema



Niveles TCP/IP

Descripción



- **Nivel físico**
 - Se encarga del acceso físico a la red
 - Gestiona las comunicaciones de la tarjeta de red, modem, o cualquier dispositivo que conecte la computadora a la red
- **Nivel de Internet**
 - Se ocupa de la transmisión de paquetes por la red
 - Encamina los paquetes por la red
- **Nivel de Transporte**
 - Se ocupa de que los paquetes se entreguen en el mismo orden en que se transmitieron sin duplicados ni pérdidas
 - Gestiona los errores
 - El protocolo TCP realiza estas tareas
- **Nivel de Aplicación**
 - Permite que los protocolos instalados en este nivel usen el canal de comunicación libres de errores
 - Ejemplos:
 - HTTP (HyperText Transfer Protocol)
 - FTP (File Transfer Protocol)

Arquitectura cliente-servidor

Protocolo HTTP (HyperText Transfer Protocol)

Navegador (Browser)

GET http://www.servidor.com/index.html



Servidor web

Escucha las peticiones en un puerto (habitualmente el 80) y responde enviando

http:/1.1 200 OK

<html>

<body>

Enlace a

Otro

</body>

</html>

Protocolo HTTP

HyperText Transfer Protocol

- Es un protocolo cliente-servidor que se usa para intercambio de información entre los clientes y servidores de la Web
- Funciona sobre los servicios de red TCP/IP
- Los servidores Web esperan las peticiones de los clientes escuchando un puerto (habitualmente el puerto 80)
- Un **puerto** es un número que identifica a una aplicación que está preparada para intervenir en una comunicación TCP.
 - La numeración se corresponde con un estándar que asocia a cada número un servicio.

Protocolo HTTP

URL (*Uniform Resource Locator*)

- Las URL especifican las direcciones de los recursos Web
- Ejemplo:
 - <http://www.w3.org>
- También pueden especificar un documento
 - <http://www.servidor.com/index.html>
 - <http://www.servidor.com/documento.pdf>

Mandatos básicos del protocolo HTTP

- **GET**. Solicita leer una página web
- **HEAD**. Solicita leer la cabecera de una página web
- **PUT**. Solicita almacenar una página web
- **POST**. Envía datos a una aplicación web
- **DELETE**. Borra la página web
- **LINK**. Conecta a dos recursos existentes
- **UNLINK**. Rompe una conexión existente entre dos recursos

Cientes Web

Navegadores (browsers)

- Los clientes se conectan al servidor y le envían mensajes a su puerto 80 usando TCP/IP
- Son programas que permiten acceder a la Web y visualizar en modo gráfico documentos HTML (XHTML)
- También pueden visualizar otros tipos de archivo como los formatos gráficos: GIF, JPG y PNG.
- Además se han añadido mecanismos para ejecutar programas a través de la red Internet
- También pueden arrancar aplicaciones que muestren los archivos recibidos
 - Por ejemplo documentos de MS-Word
- Se les puede añadir módulos denominados “**plug-ins**” para mostrar ciertos tipos de documentos
 - Por ejemplo documentos PDF (Portable Document Format)
 - Presentaciones Flash

Cientes Web

Ejemplos de Navegadores (browsers)

- El primero fue el Mosaic (1993)
- Netscape Navigator
- Internet Explorer de Microsoft
- Opera
- FireFox
- Otros

Servidores Web

Modelo de funcionamiento

- Permanecen a la escucha en un puerto (habitualmente el 80)
- Reciben por medio de mensajes las peticiones de los clientes
- Los mensajes tienen un formato bien definido en el protocolo HTTP, siendo fundamental el mandato adjunto a los mismos
- Tras procesar la operación el servidor responde con otro mensaje que puede incluir, en función de la información solicitada:
 - Estados
 - Datos
 - Códigos de error
- Todas las operaciones pueden adjuntar objetos o recursos Web descritos por su URL. Por ejemplo:
 - Documentos HTML
 - Archivos multimedia
 - Aplicaciones CGI
- El servidor y los clientes saben como manejar estos objetos y recursos gracias a las descripciones MIME que incorporan.
 - MIME: Multipurpose Internet Mail Extensions

Servidores Web

Ejemplo de servidores Web

- **Apache**

- Open Source
- Disponible para distintos entornos. Por ejemplo:
 - Linux
 - Windows
- Soporta PHP, Java, etc.



- **Internet Information Server (IIS)**

- Específico para entornos Windows
- Soporta de forma nativa la tecnología .NET
- También puede soportar PHP, Java, etc.



Tecnologías Web



- Estándares en la Web
- Lenguajes de marcas (XHTML y XML)
- Generación dinámica de la información
- Plataformas de desarrollo
 - Plataforma Java
 - Microsoft .NET
 - Open Source
 - Gestores de contenidos

Estándares en la Web

Los iniciales



- **Hipertexto (Ted Nelson, 1965)**
 - Documentos no secuenciales, con enlaces
- **Web (Tim Berners-Lee, 1990)**
 - Red de Ordenadores conectados mediante protocolo TCP/IP + HTTP
 - TCP/IP (Trasmision Control Protocol / Internet Protocol)
 - HTTP (HyperText Transfer Protocol)
 - Protocolo de transferencia de hipertexto
 - HTML (HiperText Markup Language)
 - Lenguaje de marcas para hipertexto
 - URLs (Uniform Resource Locators)
 - Sistema de direcciones de los recursos de la web
- Desde 1995 los estándares están definidos por el **World Wide Web Consortium**
 - <http://www.w3.org> Sitio internacional
 - <http://www.w3c.es> Oficina Española

Estándares en la Web

XHTML



- Lenguaje de Marcas de HiperTexto eXtensible.
- Es una versión más estricta y limpia de HTML
- Reemplaza a HTML
- XHTML extiende HTML 4.0 combinando:
 - HTML (diseñado para mostrar datos)
 - XML (diseñado para describir datos)
- XHTML puede incluir otros lenguajes (SVG, SMIL, MathML, ...) que no incorporaba HTML

Estándares en la Web

CSS



- Cascading Style Sheets
- Hojas de Estilo en Cascada
- Es un mecanismo que describe como se va mostrar un documento
 - En pantalla
 - O como se va a imprimir
 - O como se va a pronunciar
- Se utilizan para dar estilo a documentos HTML y XML
- Separan contenido de presentación
- Cualquier cambio en el estilo de una CSS afectará a todas las páginas vinculadas a esa CSS

Estándares en la Web

XForms



- Lenguaje de etiquetado para formularios Web.
- Comprueba automáticamente los valores de los datos mientras el usuario los introduce
- Indica que ciertos campos son obligatorios y que el formulario no será enviado sin esta información
- Envía los datos de los formularios como XML
- Se integra bien con los Servicios Web

Estándares en la Web

XML (*eXtended Markup Language*)

- Lenguaje de etiquetado extensible
- Metalenguaje
- Permite diseñar lenguajes para la descripción de datos en campos específicos
- Es utilizado para intercambiar datos entre diferentes aplicaciones

Estándares en la Web

Tecnologías XML

- Son un conjunto de estándares que ofrecen servicios útiles para el manejo de XML
- **XSL**. Lenguaje extensible de hojas de estilo
 - Es capaz de transformar, ordenar y filtrar datos XML, y darles formato basándolo en sus valores.
- **XPath**. Lenguaje de rutas XML
 - Permite acceder a partes de un documento XML
- **XLink**. Lenguaje de enlace XML
 - Permite insertar elementos en documentos XML, para crear enlaces entre recursos XML
- **XPointer**. Lenguaje de Direccionamiento XML
 - Permite el acceso a la estructura interna de un documento XML
- **XQL**. Lenguaje de Consulta XML
 - Facilita la extracción de datos desde documentos XML

Estándares en la Web

Accesibilidad: WCAG



- Pautas de Accesibilidad al contenido en la Web
- Nivel "A" de Conformidad: Se han satisfecho todos los puntos de verificación de Prioridad 1;
- Nivel "Doble-A" de Conformidad: Se han satisfecho todos los puntos de verificación de Prioridad 1 y 2;
- Nivel "Triple-A" de Conformidad: Se han satisfecho todos los puntos de verificación de Prioridad 1, 2, y 3.



Estándares en la Web

Tecnologías multimedia: SVG

- Gráficos Vectoriales Escalables
- Es un vocabulario XML
- Describe gráficos vectoriales en dos dimensiones
- Tiene tres tipos de objetos
 - Formas gráficas vectoriales (líneas, curvas,...)
 - Imágenes
 - Texto
- Permite crear gráficos dinámicos e interactivos
- Es compatible con otros estándares de la Web

Estándares en la Web

Tecnologías multimedia: SMIL

- Lenguaje de Integración de Multimedia Sincronizada
- Es un lenguaje basado en XML
- Utiliza etiquetas para controlar la composición y la secuencia de los eventos en una presentación multimedia
- Todos los elementos están marcados con etiquetas
- Permite la creación de presentaciones audiovisuales interactivas
- Integra audio y video con imágenes, texto y otros medios
- Permite la interacción con el usuario

Estándares en la Web

Tecnologías multimedia: PNG

- Gráficos de Red Portátiles
- Es un formato de archivo gráfico raster
- Almacena sin pérdida de calidad
- Tiene un nivel de compresión bueno
- Está libre de patentes
- Permite manejar imágenes con
 - color indexado
 - escala de grises
 - Color verdadero
- Tiene múltiples niveles de transparencias

Estándares en la Web

Interacción multimodal

- **VoiceXML** es un lenguaje de marcas que permite interactuar escuchando comandos de voz.
- **SRGS**. Gramática de reconocimiento del habla. Utilizada por VoiceXML.
- **SSML**. Lenguaje de marcas para Síntesis del Habla basado en XML
- **EMMA**. Lenguaje de intercambio de datos en sistemas de administración de interacción multimodal
- **InkML**. Lenguaje que convierte símbolos y mensajes a texto.

Estándares en la Web

Seguridad

- Son los estándares que persiguen la necesidad de garantizar la integridad, la confidencialidad y la autenticidad de los datos que fluyen a través de la Web
- **XML Encrytion**
 - Lenguaje que asegura la confidencialidad de partes de documentos XML
 - Encripta parcialmente el documento transportado
 - Se puede aplicar a cualquier recurso Web, incluso a contenidos que no son XML
- **XML Signature**
 - Es un sistema que a través de una firma digital permite ofrecer autenticidad de los datos.
 - Con la firma digital se confirma la identidad del emisor, la autenticidad del mensaje y su integridad
- **XML Key Management**
 - Es un protocolo para distribuir y registrar la claves públicas

Estándares en la Web

Privacidad

- Son los estándares que persiguen garantizar la privacidad de los datos personales que se manejan por Internet
- **P3P** (*Platform for Privacy Preferences*)
 - Plataforma de Preferencias de Privacidad
 - Es un lenguaje estándar que controla la información personal en los sitios Web que se visitan
 - Permite desarrollar herramientas y servicios que ofrezcan a los usuarios un mayor control sobre la información personal que se maneja en Internet
 - Uno de sus objetivos es aumentar la confianza de los usuarios en el uso de Internet

Estándares en la Web

Internacionalización (I)

- La Web debe funcionar
 - En cualquier país
 - En cualquier idioma
 - En cualquier cultura
- Los sitios Web se deben diseñar de forma que se adapten automáticamente a cualquier idioma y región sin necesidad de cambios de código.
- Se utiliza **Unicode/ISO 10646** para identificar y describir los caracteres
- HTML utiliza el atributo **lang** para indicar el idioma de segmentos de contenido

lang="es"

- XML utiliza el atributo **xml:lang** para indicar el idioma de segmentos de contenido

xml:lang="en"

Estándares en la Web

Internacionalización (II)

- Es un requisito imprescindible identificar el idioma de cada página Web al inicio del documento antes del **<head>**
- También es importante marcar cualquier cambio de idioma que se produzca a lo largo del texto de una página Web
- Es importante etiquetar el conjunto de caracteres
 - En XML

```
<?xml version="1.0" encoding="UTF-8" ?>
```
 - En XHTML hay que usar la etiqueta **<meta>** dentro de **<head>** de la siguiente forma

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

Estándares en la Web

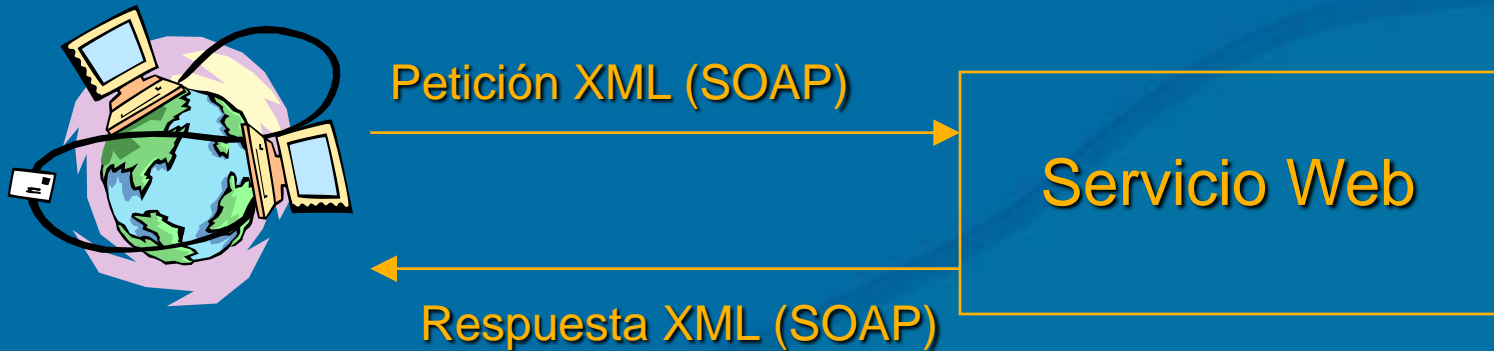
Independencia de dispositivo

- Son los estándares que persiguen que la información de la Web esté disponible y accesible en cualquier dispositivo
- Desde el punto de vista del usuario significa acceso universal
- Desde el punto de vista del desarrollador significa un único desarrollo y multitud de dispositivos
- **CC/PP**
 - Composite Capabilities/Preferences Profile
 - Sistema para expresar las capacidades de los dispositivos y las preferencias de los usuarios

Estándares en la Web

Servicios Web: Definición

- Conjunto de tecnologías y estándares que permiten interoperar automáticamente en la Web intercambiando datos y servicios entre sí.



Estándares en la Web

Servicios Web: Otra definición

- Aplicaciones modulares auto-contenidas, auto-descritas que pueden ser publicadas, localizadas e invocadas a través de la Web
- Una vez desarrollado y publicado un servicio Web, otras aplicaciones (u otros servicios Web) pueden localizarlo e invocarlo remotamente

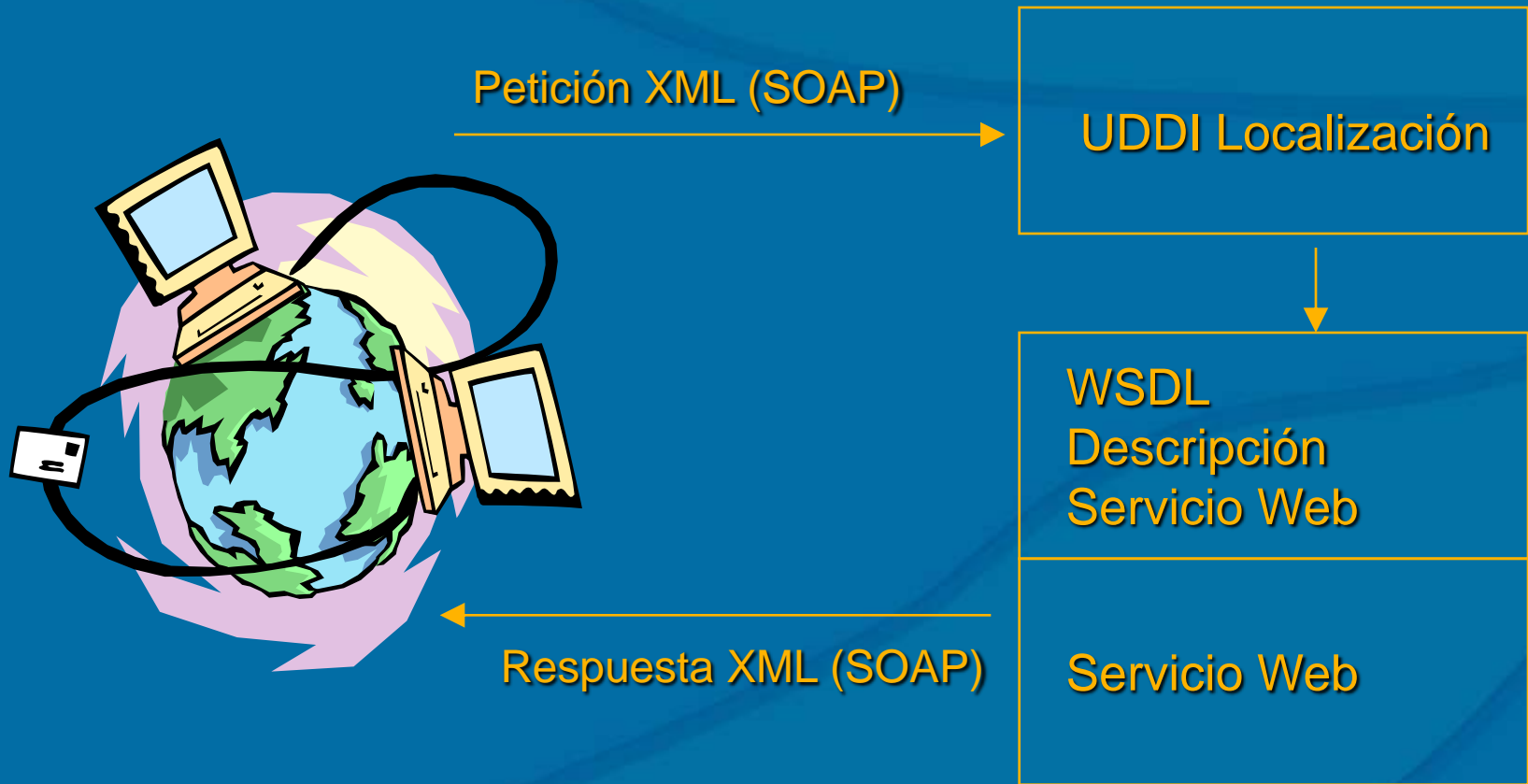
Estándares en la Web

Servicios Web: Estándares asociados

- SOAP
 - Protocolo Simple de Acceso a Objetos
- UDDI
 - Mecanismo de localización de servicios
- WSDL
 - Lenguaje de Descripción de Servicios Web
- XMLP
 - Protocolo de transferencia de archivos XML
- XKMS
 - Soporte de registro y autenticación

Estándares en la Web

Servicios Web: Funcionamiento



Estándares en la Web

La Web Semántica

- Es una Web extendida
- Basada en el significado (semántica) y no en la sintaxis
- El software debe ser capaz de procesar su contenido automáticamente
 - La Web inicial era para ser visualizada por personas, no por máquinas
- La Web Semántica como infraestructura basada en **metadatos** aporta un camino para razonar en la Web y extender sus capacidades

Estándares en la Web

La Web Semántica: Estándares (I)

- **RDF**
 - Infraestructura de Descripción de Recursos
 - Proporciona información descriptiva simple sobre los recursos que se encuentran en la Web

Estándares en la Web

La Web Semántica: Estándares (II)

- **OWL**
 - Lenguaje de Ontologías Web
 - Las ontologías son el conjunto de términos utilizados para describir y representar un área de conocimiento
 - Las ontologías incluyen la definición de conceptos básicos en un campo determinado y la relación entre ellos

Estándares en la Web

La Web Semántica: Aplicaciones (I)

- **RSS**
 - Vocabulario RDF basado en XML
 - Permite la catalogación de información
 - Noticias
 - Eventos
 - Permite encontrar información precisa adaptada a las preferencias de los usuarios

Estándares en la Web

La Web Semántica: Aplicaciones (II)

- **FOAF**
 - Proyecto de Web Semántica
 - Vocabulario RDF que contiene información personal
 - Describe
 - Personas
 - Vínculos entre personas
 - Creaciones de las personas
 - La información puede ser procesada, compartida y reutilizada
 - Los datos personales pueden utilizarse en los motores de búsqueda
 - **FOAF-a-Matic**
 - Aplicación Javascript que permite crear una descripción **FOAF** de uno mismo
 - **FOAFNAUT**
 - Muestra relaciones entre FOAF y SVG

Referencias Generales

- <http://www.di.uniovi.es/~cueva/investigacion/lineas/web>
 - Página Web del autor con enlaces a distintos temas de Ingeniería Web
- Guías breves de tecnologías W3C.
 - Oficina Española W3C
 - www.w3c.es
- Programación de Aplicaciones Web
 - S. Rodríguez de la Fuente et al.
 - Editorial Thomson, Madrid, 2003