

13  
24.



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES  
"CAMPUS ARAGON"**

**"CREACION DE PAGINAS  
HIPERMEDIA EN WWW"**

**T E S I S**  
**QUE PARA OBTENER EL TITULO DE**  
**I N G E N I E R O E N**  
**C O M P U T A C I O N**  
**P R E S E N T A :**  
**ULISES CRUZ CORTES**

**ASESOR: ING. ARMANDO REYES GONZALEZ**

**MÉXICO.**

**1997**

**TESIS CON  
FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Agradezco profundamente a la Universidad Nacional Autónoma de México, darme la oportunidad de ser parte de su comunidad y hacer posible mi formación profesional.**

**A mis papás Mateo Cruz Ángeles y Sofía Cortés de Cruz, por ser el mejor ejemplo de mi vida, gracias por contribuir en mi educación y formación personal al darme una vida llena de felicidad a su lado, por todo su apoyo y confianza a lo largo de mi existencia, papás los amo. César, gracias por compartir conmigo los años más felices de mi vida al lado de nuestros padres, agradezco tus consejos y la ayuda incondicional que siempre me has brindado, te quiero mucho. A mi abuelita Doña Rosa Cortés Martínez, por sus innumerables sacrificios y amor eterno.**

**Deseo extender un agradecimiento especial al Ingeniero Armando Reyes González Director de esta Tesis, por su confianza y apoyo incondicional, así como a la DGSCA, institución sin la cual no hubiese sido posible el desarrollo de este trabajo, a los buenos profesores que me formaron, a los verdaderos amigos.**

**Sinceramente**

**Ulises.**

***A la memoria entrañable de mis abuelos  
Don Agustín Cruz Morales y Trinidad  
Ángeles de Cruz, con cariño y respeto por  
sus innumerables sacrificios a lo largo de  
sus vidas. . . .Nunca los olvidaré.***

# **“Creación de páginas hipermedia en WWW”**

---

*Ulises Cruz Cortés*

## Objetivos

---

- Exponer cuales son los principios y elementos básicos que intervienen en el proceso de creación de páginas para "World Wide Web", a través del lenguaje HTML y sus extensiones.
  - Proporcionar mediante la información contenida en el presente trabajo una guía, que sirva como base al desarrollo de aplicaciones más detalladas para Internet en idioma español.
  - Destacar el impacto actual del Web como medio de información y comunicación *global sin fronteras*, así como el gran crecimiento que presenta.
-

# Contenido

---

<b>INTRODUCCIÓN</b>	<b>1</b>
<b>CAPITULO I</b>	
<b>INTERNET</b>	<b>3</b>
1. Historia	3
2. Concepto de red universal	10
3. Arquitectura de Internet	10
4. Direccionamiento	16
5. Modelo cliente servidor	19
6. Aplicaciones y servicios	20
<b>CAPITULO II</b>	
<b>WORLD WIDE WEB</b>	<b>25</b>
1. HTTP	26
2. Componentes de WWW	31
3. Hipertexto e Hipermedia	32
4. Servidores y clientes en WWW	33
<b>CAPITULO III</b>	
<b>Lenguaje HTML</b>	<b>37</b>
1. Etiquetas HTML	38
2. Formato del URL	40
3. Estructura de un documento	43
4. Encabezado	43
5. Cuerpo	45
5.1 Espaciados y saltos de línea	47
5.2 Caracteres especiales	49
5.3 Encabezados	50
5.4 Tamaño y color de caracteres	51
5.5 Estilos físicos y lógicos	53

5.6	Listas	54
5.7	Imágenes	56
5.8	Hiperenlaces	58
5.9	Tablas	60

**CAPITULO IV  
EXTENSIONES HTML 65**

1.	Formas	65
2.	Common Gateway Interface	69
3.	Frames	73
4.	Etiquetas Especiales	78
4.1	Applet	78
4.2	Marquee	79
4.3	Elementos multimedia	80
4.4	Extensiones Internet Explorer	81
4.5	Extensiones Netscape	82
4.6	Scripts	82

**CAPITULO V  
CASO DE APLICACIÓN 85**

Sitio Web del Departamento de Productos Interactivos para la Docencia DGSCA	85
-----------------------------------------------------------------------------------	----

**CONCLUSIONES 95**

**APÉNDICES:**

<b>A.-</b> Metodología de Desarrollo	99
Guía de Estilo	104
<b>B.-</b> Proveedores de Acceso en México Organismos Internacionales	109
<b>C.-</b> Estándares de Comunicación	113

<b>GLOSARIO</b>	117
-----------------	-----

<b>REFERENCIAS</b>	123
--------------------	-----

# Introducción

---

**Hasta** hace algunos años, nadie hubiera imaginado el desarrollo que ha alcanzado la industria de la computación, inicialmente calificada como una tecnología complicada y de uso limitado a científicos e investigadores, en menos de 45 años ha logrado crecer a tal grado que se ha vuelto indispensable para desempeñar actividades comunes y cotidianas en la sociedad; sin lugar a duda, se ha convertido en el elemento central de la revolución tecnológica actual, ya que a logrado incrementar de manera inimaginable la productividad a un ritmo impresionante, constituyéndose en una valiosa herramienta para lograr una aceleración en los diferentes campos del saber humano.

Gracias a los avances tecnológicos en materia de telecomunicaciones como son: enlaces satelitales, computación remota, telefonía celular, televisión directa al hogar, etc., se ha facilitado la conectividad mundial, permitiendo una más ágil y oportuna comunicación entre los seres humanos y las diversas actividades que realizamos, ya sean de carácter: económico, político, social o cultural. La simplificación en el tamaño de los equipos, así como el gran poder de procesamiento y facilidad de uso, caracterizan estas tecnologías de fin de siglo.

Internet surge de estos avances, constituyéndose en una gigantesca red de computadoras distribuida en todo el mundo, en la que se puede encontrar información, gente y servicios de todo tipo; *el éxito de Internet radica en la libertad que ofrece*, no existe un control central o manipulación de algún tipo, no hay censura, ni limitantes.

Una de las aplicaciones más importantes en Internet es World Wide Web (mejor conocido como WWW), o simplemente Web, tema central de esta Tesis ya que *actualmente se trata del medio funcional para que la gente de todo el mundo localice información de cualquier tipo y comparta en forma general el conocimiento humano*, razón suficiente que me motivó a profundizar en su estudio y campos de aplicación, ya que además de ser el impulsor de la enorme difusión de Internet, se trata del medio de comunicación que está transformando a nuestra sociedad, con un impacto mayor al de la televisión o cine, aunado al enorme crecimiento que presenta y a los avances tecnológicos de cada mes, *uno puede afirmar con certeza que WWW transformará por completo la forma de vida de la humanidad*.

La presente Tesis muestra los principios y elementos básicos que intervienen en el proceso de creación de páginas hipertexto (base funcional de WWW), a través del lenguaje HTML; para brindar un marco referencial y ubicar el extenso tema que implica

## *Introducción*

hablar de Internet, expongo en el **Capítulo I**, su historia, arquitectura y funcionamiento, también las aplicaciones y servicios más comunes que se pueden encontrar en esta gran red. El **Capítulo II** describe extensamente a World Wide Web, los componentes y servicios que soporta, el protocolo HTTP del cual depende, así como una breve semblanza de servidores y clientes propios de esta tecnología. Los **Capítulos III y IV** son en esencia una exposición del lenguaje HTML y sus extensiones, estos capítulos proporcionan una guía de referencia ilustrada con ejemplos, que cubren a la gran mayoría de las etiquetas y atributos propios de este lenguaje de marcación de hipertexto, el cual da formato a páginas del Web. Finalmente en el **Capítulo V** expongo el caso de aplicación particular de este trabajo de Tesis, en el se ubican las necesidades del departamento de Productos Interactivos para la Docencia de la DGSCA (UNAM), así como la solución a la problemática encontrada, también describo el proceso de análisis, diseño, desarrollo e implementación del Sitio Web del departamento y su estado actual; al final de este capítulo proporciono las **conclusiones** más relevantes y satisfactorias propias de la aplicación, encontradas a lo largo de casi 8 meses de investigación documental y más de un año de desarrollo en Internet, a través de la tecnología World Wide Web.

Es conveniente para el lector que no está familiarizado con cierta terminología de computación, redes e Internet, tener presente la ubicación al final de esta tesis de un **glosario de términos**, el cual resulta de gran valor en conjunto con las referencias a otras fuentes que proporciono.

Gracias a la información contenida en el presente trabajo, proporciono una guía práctica que sirve como base para el desarrollo de aplicaciones más avanzadas y detalladas, no solo aplicables al área de Ingeniería, sino a cualquier otra, gracias al carácter multidisciplinario y flexibilidad de WWW.

# Capítulo I

## Internet

---

Internet es una gigantesca "red de redes", distribuida en todo el mundo, la cual se comunica a gran velocidad por medio del conjunto de protocolos TCP/IP<sup>1</sup> y que funciona como una sola *red virtual global cooperativa*, proporcionando una conectividad transparente entre plataformas heterogéneas y sistemas, en la que se puede acceder a información de todo tipo, a través de ciertos servicios y herramientas tales como: Archie, Wais, Gopher, Mail, **World Wide Web** o mejor conocido como WWW, de las cuales se hablará más adelante en el capítulo.

No existe una definición precisa que englobe a todo lo que comprende Internet; puede ser definida: *en relación a sus protocolos comunes*, como una *colección física de ruteadores, gateways y circuitos*, o un *conjunto de recursos compartidos*, y hasta como una *actitud acerca de interconexión e intercomunicación mundial*; lo cierto es que actualmente Internet, es la red de computadoras más grande del mundo, que permite una comunicación global entre empresas, institutos, gobiernos, universidades, centros de trabajo, corporaciones, milicia, etc.

La importancia de esta enorme red radica, además de los servicios usuales que se pueden obtener de ella, en el interés general que las personas tienen, pero sobre todo, en la comunicación global sin fronteras ni censura que proporciona, e información libre entre personas y grupos sociales alrededor del mundo.

### 1.1 Historia

Internet fue creado a partir de un proyecto del Departamento de Defensa de los Estados Unidos de Norteamérica (EUA), llamado **DARPA**NET (Defense Advanced Research Project Agency Network), durante la guerra fría (después de la Segunda Guerra Mundial) y cuyo propósito principal era la investigación y desarrollo de protocolos de comunicación

---

<sup>1</sup> Transmission Control Protocol/Internet Protocol.

para redes de área amplia, para poder ligar redes de transmisión de datos de diferentes tipos, capaces de resistir condiciones de operación difíciles y continuar funcionando normalmente, aún con el daño parcial o pérdida de una parte de la red (por ejemplo, en caso de guerra nuclear).

Los principios fueron simples: todos los nodos de la red deberían ser iguales —*misma jerarquía*—, cada uno con autonomía suficiente para originar, recibir y retransmitir mensajes, además cada mensaje debería ser dividido en paquetes de información y cada uno de ellos direccionado<sup>2</sup>, así la ruta tomada por el paquete, para llegar a su destino no tenía que ser importante.

A principios de 1969 el primero de tales nodos fue instalado en la UCLA (Universidad de California en Los Angeles) y para diciembre de ese año había ya cuatro nodos en la red. Durante los 70's DARPA/NET creció: su *estructura descentralizada* facilitó su expansión y, a diferencia de otras redes, pudo unir diferentes tipos de computadoras. El protocolo original fue NCP (Network Control Protocol), pero al paso del tiempo y avance de la tecnología, éste fue sucedido por un estándar más sofisticado. Debido a su crecimiento, a finales de 1983 se dividió en dos subredes principales: ARPANET y MILNET, esta última dedicada a conectar centros militares de cómputo con bases estratégicas distribuidas en territorio de los EUA.

El crecimiento de las computadoras conectadas a ARPANET fue vertiginoso, entonces, surgió la necesidad de tener una normalización en los protocolos de comunicación para prevenir problemas futuros. Las investigaciones dieron como resultado el conjunto de protocolos TCP/IP (Transmission Control Protocol / Internet Protocol), un sistema de protocolos de comunicación sólido y robusto, bajo el cual se integran todas las redes que conforman lo que se conoce actualmente como Internet.

Durante el desarrollo de este conjunto de protocolos, se incrementó notablemente el número de redes de computadoras locales, de agencias gubernamentales, universidades y centros de investigación tanto en los EUA, como fuera de éste, los cuales participaban directamente en el proyecto, dando origen así a la red de computadoras más grande del mundo. Una vez separadas las funciones militares, se permitió el acceso a todo aquel que así lo solicitara, siempre y cuando se empleara el acceso para fines académicos, científicos o de investigación (y por supuesto que pagaran sus propios gastos de conexión y equipo necesarios); los usuarios con acceso a la red pronto encontraron que la información en ésta era de gran utilidad y si cada quien aportaba algo, se enriquecía aún más el acervo existente.

Por extraño que parezca NO existe una autoridad central que controle el funcionamiento de Internet, aunque existen organizaciones y grupos que se dedican a *organizar de alguna forma*, las diversas actividades y proyectos que se llevan a cabo (ver Apéndice B, Organismos Internacionales); la National Science Foundation, fue una de éstas organizaciones que promovió extensamente el uso de la red, se encargó de conectar cinco centros de supercómputo alrededor del mundo, los cuales podían ser comunicados desde cualquier nodo en la red, esto funcionó al principio, sin embargo pronto fueron superadas las cargas de tráfico previstas, entonces se dio la concesión a Merit Network

---

<sup>2</sup> Es decir, cada paquete contendría información acerca del nodo fuente y nodo destino.

Inc., para que administrara y actualizara la red, así se mejoraron las líneas de comunicación dando un servicio mucho más rápido; sin embargo este proceso de mejora nunca termina, debido a la creciente demanda (servicios) que exige la comunidad en general.

El grupo de mayor autoridad, debido al conjunto de personalidades que agrupa en el desarrollo de la red es Internet Society (ISOC), creado en 1990 y formado por miembros voluntarios, cuyo propósito principal es promover el intercambio de información mundial a través de esta tecnología, se trata de una sociedad —al estilo de un consejo—, que tiene la responsabilidad de la administración técnica y coordinación en Internet, sin embargo no es el único grupo que puede tomar decisiones, existen otros tres grupos que tienen un papel relevante (NIC, IAB y IETF).

Internet Architecture Board (IAB), toma decisiones acerca de estándares de comunicación entre las diferentes plataformas, para que puedan interactuar máquinas de diferentes fabricantes y plataformas sin problemas, este grupo es responsable de cómo asignar las direcciones y otros recursos en la red, aunque no son ellos quienes se encargan de hacer estas asignaciones, para eso existe otra organización llamada NIC (Network Information Center), administrado por el departamento de defensa de los EUA. El Internet Engineering Task Force (IETF), es otra organización en donde usuarios reales expresan opiniones sobre cómo se deben implementar soluciones para problemas operacionales y cómo deben cooperar las diversas redes para lograrlo.

Parte del enorme crecimiento de la red global, se debe a que inicialmente se encontraba basada en fondos gubernamentales de cada país que formaba parte de Internet, lo que propiciaba un servicio prácticamente gratuito (subsidiado). A principios de 1994 comenzó un crecimiento explosivo de compañías con fines comerciales, dando así origen a una nueva etapa en el desarrollo de la red. Internet NO es una compañía que cobre a todas las redes o usuarios que la integran, cada institución paga por su conexión a alguna red regional, la cual a su vez paga acceso a algún proveedor nacional o internacional.

**Cronológicamente**, la historia de Internet puede ser observada en la siguiente tabla:

Año	Suceso
1957	La Agencia de Proyectos de Investigación Avanzados del Departamento de Defensa de los EUA es creada con la finalidad de establecer liderazgo en ciencia y tecnología aplicable a objetivos militares.
1964	Surge la necesidad por parte de los EUA de resolver un problema de estrategia militar, ¿Cómo comunicar a las autoridades gubernamentales después de una guerra nuclear?  RAND Corporation, una de las empresas encargadas de la estrategia militar estadounidense propuso una solución: la creación de una red de comunicaciones que no dependiera de un organismo central, integrado por nodos o puntos de enlace de igual rango y con la misma capacidad de originar, transmitir y recibir mensajes, y que en caso de que alguno de estos nodos recibiera un ataque o dejara de funcionar, el resto de la red seguiría en operación. Los mensajes en esta red se dividirían en paquetes de información, cada uno con su propia dirección, originado en algún nodo en particular y saltando de lado a lado hasta finalizar en otro nodo específico de manera individual. La ruta de los paquetes no importaba, lo primordial era llegar al destino.

*Creación de Páginas en WWW*

Año	Suceso
1967	La planeación de este tipo de redes se expuso durante el simposium realizado en Inglaterra sobre Principios Operativos, auspiciada por ACM (Association of Computer Machinery).
1968	El primer resultado en este tipo de redes se obtuvo en Gran Bretaña, utilizando un mainframe IBM.
1969	<p>Advanced Research Projects Agency (agencia del Pentágono), decide realizar un proyecto mayor sobre esta tecnología en redes. El proyecto fue desarrollado por: RAND Corporation, el MIT (Massachusetts Institute of Technology) y la UCLA en donde fue instalado. Para diciembre de ese año ya existían cuatro nodos en ARPANET, pudiendo transmitir datos en líneas de alta velocidad y programar remotamente computadoras en otros nodos. En 1971 había quince nodos y para 1972 treinta y siete.</p> <p>Poco a poco comenzó a expandirse el uso de ARPANET, no sólo se dedicaba a trabajos de cómputo remoto, sino que extendió la comunicación internacional de proyectos y trabajos entre científicos e investigadores, y permitió el uso personalizado de correo electrónico. Así surge además de la comunicación personal, las listas de interés, que son mensajes de correo electrónico retransmitidos automáticamente a los suscriptores en la lista de discusión.</p>
1972	Se crea INWG (Internetworking Working Group) con el objetivo de establecer protocolos comunes.
1973	Tuvo lugar la Primera Conferencia Internacional de Comunicaciones entre Computadoras (presentando a ARPANET), con la demostración de 40 computadoras conectadas entre sí alrededor del mundo, y sin ninguna pérdida de información, logrando un gran impacto. Otra ventaja de ARPANET es que no importaban los tipos o tamaños de las computadoras en las que se estuviera trabajando, mientras cumplieran con los protocolos establecidos, estas funcionarían dentro de la red.
1974	<p>El protocolo original conocido como NCP fue cambiado por un nuevo estándar más sofisticado, llamado TCP/IP, publicado en este año por Vint Cerf y Bob Kahn.</p> <p>TCP convierte mensajes en cadenas de paquetes en el nodo origen, y los ensambla de nuevo en el punto destino. IP maneja el direccionamiento permitiendo que los paquetes sean ruteados a través de diferentes nodos y hasta de diferentes redes con varios estándares, como Ethernet, FDDI y X.25.</p>
1976	Se desarrolla la tecnología UUCP (Unix to Unix CoPy) en Laboratorios Bell de AT&T.
1977	Comienza a extenderse el uso de TCP/IP en otras redes para vincularse a ARPANET, comenzando esta red a volverse más pequeña en comparación con la gran cantidad de computadoras que se conectaban. A fines de los 70's y en los 80's, personas de diferentes grupos sociales tienen acceso a computadoras de gran capacidad, siendo relativamente sencillo conectarse a la creciente "red de redes". Como el software de TCP/IP es de dominio público, y por su naturaleza descentralizada, comienza el auge de la conexión a Internet. En esta época surge USENET, el boletín electrónico más grande del mundo basado en UUCP (UNIX al paso de los años se ha convertido en el sistema operativo estándar de todas las computadoras de mediano y gran tamaño conectadas a Internet). También surgieron servicios enfocados a la diversión como el primer MUD (Multi User Dungeon), juego de rol interactivo.

Año	Suceso
1981	<p>Surge BITNET (Because It's Time for NETwork), creado como red cooperativa, que provee a sus usuarios de correo electrónico, listas de discusión así como de transferencia de información.</p> <p>Se forma el CREN (Corporation for Research and Education Networking) a partir de la unión de CSNET y BITNET.</p> <p>La conexión a Internet tiene un mínimo costo, ya que cada nodo es independiente, y maneja por sí mismo sus propias necesidades técnicas y financieras. De esta manera, la red comenzó a extenderse, abarcando un mayor número de personas conectadas y de recursos disponibles. Así, la comunicación a través de la computadora comienza a ser indispensable.</p>
1982	<p>El Departamento de Defensa de los EUA declara como estándar al conjunto de protocolos TCP/IP, separándose de ARPANET la parte militar.</p> <p>Se da el auge por las estaciones de trabajo de escritorio, con sistema operativo "UNIX Berkeley" (desarrollado en la Universidad de Berkeley), que incluye software de conectividad a red con TCP/IP.</p> <p>Se crea la EUNET (Red Europea), para proporcionar servicio de correo electrónico.</p>
1983	<p>Se desarrolla el Servidor de Nombres, de esta manera no se tiene que memorizar direcciones completas de otros sistemas.</p> <p>Se establece la red EARN (European Academic and Research Network) y la IAB (Internet Activities Board), la primera con el propósito y funciones similares a Bitnet y la última para vigilar el desarrollo de Internet.</p>
1984	<p>NSF (National Science Foundation), a través de su Oficina de Cómputo Científico Avanzado establece un nuevo avance técnico, al integrar 5 supercomputadoras con enlaces más rápidos, impulsando así el desarrollo de Internet, y permitiendo una mayor cantidad de conexiones, principalmente universidades, con fines académicos y de investigación. También surge el primer Freenet (acceso público a correo electrónico y servicios de Internet en forma gratuita) en Cleveland.</p> <p>Se desarrolla DNS (Domain Name Server), en este punto se inició la organización de los dominios (o direcciones de Internet para las diferentes redes conectadas) por su ubicación geográfica, y los seis básicos: gov, mil, edu, com, org y net, que corresponden a instituciones gubernamentales, militares, educacionales, comerciales, no comerciales, y destinados a enlaces entre redes, respectivamente.</p>
1986	<p>Se diseña el protocolo NNTP (Network News Transfer Protocol) para la diseminación de noticias.</p> <p>Se desarrollan los registros MX (Mail Exchanger) para el intercambio de mensajes entre redes que no poseen conexión IP.</p>
1987	<p>El número de hosts conectados a Internet rebasa los 10,000.</p>
1988	<p>Empiezan a surgir problemas en la red, como el caso del virus "gusano" (Internet Worm), que aprovechaba un error en el código de los programas de correo electrónico.</p>

Año	Suceso
	<p>afectando a 6,000 de las 60,000 computadoras conectadas a Internet. Por este motivo DARPA crea el CERT (Computer Emergency Response Team), que genera recomendaciones y alertas en caso de problemas dentro de la Red.</p> <p>La comunicación personal tiene mayores posibilidades con el desarrollo de IRC (International Relay Chat), que permite la conversación simultánea de varias personas en todo el mundo conectadas a esta red.</p>
1989	<p>México ingresa a Internet a través de NSFNET, contando además con la red BITNET, esto permite que usuarios del ITESM y la UNAM tengan acceso a los recursos existentes en Estados Unidos de Norteamérica y el resto del mundo.</p> <p>Como ironía y muestra de la eficiencia del sistema, durante la guerra del Golfo Pérsico en este mismo año, los ejércitos de Irak utilizan Internet como medio de comunicación para sus operaciones y ataques. No pueden ser detectados por las fuerzas militares estadounidenses, ya que los iraquíes utilizan direcciones piratas, y cambian constantemente de lugar el equipo y las instalaciones.</p> <p>La importancia de Internet comienza a revelarse, al ser el único medio de comunicación sin censura ni restricciones que poseen los estudiantes chinos que se amotinan, pidiendo democracia en su gobierno opresor. También juega un factor de peso en el intento de "golpe de Estado" realizado en la Unión Soviética, ya que algunos moscovitas poseían enlace a Internet y conseguían de primera mano la información necesaria sobre la rebelión para difundirla a nivel internacional.</p>
1990	<p>ARPANET se vuelve obsoleta y deja de existir. Por iniciativa de los usuarios, surgen las primeras organizaciones dedicadas a la protección de los derechos de las personas conectadas a Internet, este es el caso de EFF (Electronic Frontier Foundation), y la primera organización que comercializa el acceso a Internet vía módem: The World.</p> <p>Se implementan herramientas que catalogan y facilitan el acceso a Internet: Archie, para la búsqueda de archivos accesibles mediante FTP (File Transfer Protocol); Hynetel, un catálogo de recursos y bibliotecas en línea accesibles mediante TELNET (terminal remota); WAIS (Wide Area Information Service), para entregar directamente documentos al usuario; Gopher, para ver la información a través de menús; PGP (Pretty Good Privacy), para dar seguridad y privacidad a los mensajes de la comunidad en la red; Veronica, un sistema de búsqueda complementario a Gopher.</p>
1991	<p>Commercial Internet eXchange (CIX) Association, Inc., surge a partir de que NSF levanta las restricciones que existían para el uso comercial de la Red.</p>
1992	<p>Es un año de profundos cambios dentro de Internet. Se funda Internet Society (ISOC), para coordinar el uso de las tecnologías existentes en beneficio de los usuarios. Se desarrolla en el CERN la tecnología World Wide Web, que permite un acercamiento más fácil a través de hipertexto a los recursos en Internet; también se da la primera muestra de audio y video en tiempo real a través de la Red.</p>
1993	<p>InterNIC es creado por NSF para proveer servicios de información, así como registros, directorios y bases de datos referentes a Internet.</p> <p>El Presidente de los EUA, Bill Clinton, su esposa Hillary y el vicepresidente Al Gore ingresan al WWW mediante la página de la Casa Blanca. En este momento los medios masivos de comunicación tradicionales (televisión, radio, cine, revistas y publicaciones</p>

Año	Suceso
	<p>en general) toman conciencia de Internet y sus implicaciones, es entonces que empiezan a aparecer artículos en revistas como: Time, Newsweek y Fortune, además de reportajes en las cadenas más importantes de televisión.</p> <p>El crecimiento de la red se vuelve exponencial gracias a WWW. Mosaic (explorador gráfico de Internet), es desarrollado en la Universidad de Illinois, la primera en aprovechar la gran capacidad del WWW, teniendo un crecimiento en ese año de 341.634% de usuarios de esta herramienta.</p>
1994	<p>Internet cumple 25 años de servicio, el Senado de los EUA provee información por este medio y los centros comerciales llegan a la red, tal es el caso de Internet Shop Network y JCPenny. El auge es tal que surgen servicios bancarios en la red, como First Virtual y los negocios comienzan a prosperar, como es el caso de Pizza Hut.</p> <p>Surge el caso de Canter &amp; Siegel, que, sin respetar las reglas de cortesía de la red (conocidas como <i>netiquete</i>), inundan USENET con anuncios sobre sus servicios para inmigración, teniendo una respuesta hostil por parte de los ciudadanos de la red (<i>net.citizens</i>).</p>
1995	<p>Los sistemas de servicios vía módem (Compuserve, Prodigy, Genie) comienzan a ofrecer servicios de Internet. Gran cantidad de compañías relacionadas con la red se vuelven privadas, encabezadas por Netscape (líder en el mercado de WWW), que tiene el tercer índice de ganancias jamás conseguido en Wall Street, el primer índice lo posee Microsoft.</p>
1996	<p>Surgen diversas alianzas comerciales entre compañías líderes en el mercado de software para WWW e Internet, encabezadas principalmente por Sun, Netscape, Intel y Microsoft.</p>

Tabla 1.1a. Historia cronológica de Internet

Durante algunos años fue limitado el uso comercial de Internet, debido al sustento casi total por parte de fondos gubernamentales y a su carácter académico y científico. Las políticas que restringían el uso de la red han cambiado, lo cual hoy se traduce en beneficios para las pequeñas compañías que no cuentan con los recursos necesarios para mantener una red de comunicaciones. Gradualmente se han ido eliminando las políticas restrictivas, permitiendo así dar un impulso a las diversas industrias que aprovechan esta tecnología, además de incrementar los servicios que se ofrecen en la red, se han reducido los costos y el usuario común puede adquirir este tipo de servicios a precios accesibles.

Las tecnologías se han mejorado para adaptarse a la enorme demanda que se espera en los próximos meses, el incremento en la velocidad de transmisión y la reducción en los costos amplían la gama de servicios existentes, los cuales permitirán muy pronto servicios que hoy no son viables. Al mismo tiempo que ocurre la comercialización se da la privatización. Durante años la comunidad Internet ha solicitado a compañías telefónicas que provean conexiones IP<sup>3</sup> de la misma forma que proporcionan líneas telefónicas, ahora que estas comienzan a descubrir el gran potencial de Internet, ven un gran mercado lleno de oportunidades e inversión.

<sup>3</sup> Conexiones basadas en el protocolo de comunicación: Internet Protocol.

## 1.2 Concepto de red universal

Una interconexión de miles de redes como es el caso de Internet, tiene como propósito principal permitir una comunicación entre sus nodos (hosts), sin importar en que red se encuentre localizado, de manera transparente al usuario. La ideología de Internet es muy flexible: *"si se desea conectar una nueva red, simplemente hay que establecer una conexión con algún punto de Internet".*<sup>4</sup>

El concepto de red universal se basa en un esquema de redes interconectadas globalmente, el cual permite tener un sistema unificado de redes que soportan un servicio de comunicación, la ventaja radica en que el diseño conceptual oculta la arquitectura de red al usuario, este último NO necesita conocer detalles de hardware y protocolos de comunicación afines. Así es posible comunicar datos entre dos nodos NO conectados directamente (mediante redes intermedias), gracias al conjunto de protocolos TCP/IP y a que todas las computadoras tienen asignado un identificador único y universal, además la interfaz de los programas o aplicaciones finales con el usuario son independientes del hardware en la red.

## 1.3 Arquitectura de Internet

La arquitectura o estructura organizacional de Internet, se refiere a la interconexión física de redes que la conforman, la cual se realiza mediante computadoras, o equipos de procesamiento de datos con dispositivos de comunicación llamados "gateways" (computetas), pudiendo recibir otros nombres como ruteadores o puentes. Estos dispositivos cuentan con conexiones directas a otras redes y son los encargados de transferir paquetes de información, observe la figura 1.3a. La función de un gateway consiste en identificar paquetes de información destinados a nodos que NO se encuentran en la misma red en la cual se originaron, una vez que son identificados y capturados, los paquetes serán transmitidos a la red correspondiente.



Las redes se comunican mediante compuertas, todas las redes vistas como iguales.  
Figura 1.3a **Arquitectura global de interconexión de Redes en TCP/IP.**

<sup>4</sup> Comer, Douglas *"Internetworking with TCP/IP", vol. 1*  
1995 Prentice Hall, pag 56.

Para poder actuar como conexiones entre redes, los gateways (G) deben realizar funciones de ruteo, es decir, elegir el conjunto de redes intermedias para que un paquete de información llegue a su destino correcto. Para lograr esto, cada gateway debe conocer la topología de Internet más allá de las redes a las que se encuentra conectado directamente. La información correspondiente a la topología, se mantiene en tablas almacenadas en cada uno de los gateways de Internet. Así el usuario ve la interconexión de redes como si se tratara de una sola red virtual (transparentemente), que le permite establecer un canal de comunicación entre computadoras de diferentes plataformas, sistemas operativos, incluso si éstas son de arquitectura de hardware diferente.

Es importante tener en mente que *los protocolos de comunicación establecen una descripción formal de los formatos que deberán presentar los mensajes para poder ser intercambiados por equipos de cómputo, además definen las reglas que ellos deben seguir para lograrlo.* Así el conjunto de protocolos TCP/IP, identifica a todas las redes de la misma manera, sin importar si estas son de área local (LAN) o área amplia (WAN).

Dentro del conjunto de estándares involucrados en Internet, TCP/IP es el más importante, debido a que es el conjunto de reglas de comunicación que se ha adoptado con base en el modelo de referencia OSI.

## Modelo OSI

La Organización Internacional de Estándares (ISO, por sus iniciales en inglés), es una organización mundial que tiene como objetivo definir una nueva arquitectura de redes estratificadas, conocida como modelo OSI (Open Systems Interconnections). Esta nueva arquitectura intenta servir como base para los nuevos protocolos que eliminarán algunas de las limitaciones inherentes a TCP/IP, tales como el máximo tamaño de un mensaje. Se espera que en los próximos 10 años, los protocolos basados en OSI, vayan reemplazando gradualmente a TCP/IP. En la actualidad el modelo de referencia OSI aún no se ha logrado implementar y constantemente es sometido a pruebas.

Este modelo se encuentra definido por varios comités internacionales, compuestos por ciertos gobiernos, universidades, institutos y corporaciones. Su objetivo es diseñar una arquitectura de red que elimine limitaciones a los protocolos existentes, así como facilitar una expansión modular, con varias funciones de red que dividen en 7 capas a este modelo:

<b>7. Aplicación:</b>	Provee servicios de alto nivel para usuarios finales, tales como: transferencia de archivos, correo electrónico o acceso a terminales remotas.
<b>6. Presentación:</b>	Realiza funciones comunes de manipulación de datos, tales como: compresión, encriptación, transacción de código ASCII/EBCDIC, etc. Efectúa cambios de formato, (en formato de red a formato host), siempre que sea necesario.

5. <b>Sesión:</b>	Controla la conexión de los procesos de comunicación, la secuencia de datos y el enlace entre un protocolo y una aplicación específica.
4. <b>Transporte:</b>	Controla la conexión punto-a-punto entre hosts. por ejemplo: el control de flujo, ajuste de velocidad de transmisión, corrección de errores y retransmisión de mensajes.
3. <b>Red:</b>	Esta capa controla el enrutamiento de información a través de la red, controla el congestionamiento de esta, así como la fragmentación y reensamble.
2. <b>Enlace:</b> (liga de datos)	Maneja el hardware de la red. Durante la manipulación de una máquina a otra, verifica el equipo y provee multiplexaje para varias líneas lógicas sobre una sola línea física.
1. <b>Físico:</b>	Es el hardware de los componentes físicos de la red (conectividad entre host).

Tabla 1.3a Capas que componen al modelo de referencia OSI.

**TCP/IP** (Transmission Control Protocol / Internet Protocol) es un conjunto de protocolos independiente del hardware, sistema operativo y tipo de red que se utilice, su esquema de direccionamiento es común a cualquier dispositivo que lo emplee, además NO es necesario tener una conexión a Internet para poderlo usar, es decir, se puede emplear este conjunto de protocolos para intercomunicar una LAN, sin necesidad de tener una conexión (como el caso en una Intranet). TCP/IP es un grupo de protocolos diseñados con una arquitectura en capas. Las capas permiten dividir en módulos las tareas y servicios que realiza, este diseño también especifica la manera en que un módulo interactúa con otros. Tomando al modelo OSI como referencia, se puede afirmar que para cada capa o nivel que él define, existen uno o más protocolos interactuando. Los protocolos son "entre-pares" (*peer-to-peer*), es decir, un protocolo en algún nivel dialoga con el protocolo del mismo nivel en la computadora remota.

En la actualidad las funciones propias de una red de computadoras pueden ser divididas en las siete capas propuestas por ISO, para su modelo de sistemas abiertos OSI. Sin embargo la implantación real de una arquitectura puede diferir de este modelo. Las arquitecturas basadas en TCP/IP proponen cuatro capas en las que las funciones de las capas de *Sesión* y *Presentación en OSI*, son responsabilidad de la capa de Aplicación y las capas de *Liga de Datos* y *Física de OSI*, son vistas como la capa de Acceso al medio. Por tal motivo para TCP/IP sólo existen las capas de: **Acceso al medio, Intercomunicación en Red, Transporte y capa de Aplicación.**

### Capa de Acceso al medio

Esta capa inferior maneja un conjunto de rutinas que permiten el acceso entre redes a nivel físico, emitiendo al medio, flujos de bits y recibiendo los que de él provienen. Define

cómo utilizar la red para transmitir un datagrama IP. Un datagrama es un formato de paquete o entidad que tiene la información suficiente para ser enrutado de una máquina origen a otra destino.

## Capa de Intercomunicación en Red

Esta capa controla la comunicación entre un equipo y otro, decide que rutas deben seguir los paquetes de información para alcanzar su destino, es decir, define el datagrama y maneja el ruteo de los datos. Conformar los paquetes IP que son enviados por la capa inferior y reencapsula los paquetes recibidos pasándolos a la capa superior.

IP (Internet Protocol), es el bloque constructor de Internet, sus funciones son las de definir el datagrama (unidad básica de transmisión), definir el esquema de direccionamiento, mover los datos entre la capa de acceso al medio (o acceso a la red) y la capa de transporte host-to-host, enrutar los datagramas a las máquinas remotas y realizar la fragmentación y reensamble de estos.

## Capa de Transporte

La capa de Transporte provee comunicación punto a punto desde un programa de aplicación a otro, regulando el flujo de información. Proveer un transporte confiable asegurándose de que los datos lleguen sin errores y en la secuencia correcta, además coordina a múltiples aplicaciones que se encuentren interactuando con la red simultáneamente de tal manera que los datos que envía una aplicación sean recibidos correctamente por la aplicación remota, esto lo hace añadiendo identificadores de cada una de las aplicaciones. Realiza además una verificación por suma, para asegurar que la información no sufrió alteraciones durante su transmisión.

Esta capa contiene dos importantes protocolos: el Protocolo de Control de Transmisión (TCP), el cual provee el servicio de entrega de datos con una detección y corrección de errores end-to-end; y el Protocolo de Datagrama de Usuario (UDP), que es un servicio orientado a no-conexión, es decir que NO posee técnicas para verificar si los datos llegaron a su destino correctamente.

## Capa de Aplicación

Se trata de la capa más alta del modelo, esta provee servicios de alto nivel a usuarios, invoca programas que acceden a servicios en la red, e interactúa con uno o más protocolos de transporte dependiendo del tipo de servicio que requieran para enviar o recibir información, esta capa la conforman las aplicaciones y procesos que comúnmente son utilizados, de los cuales se hablará más adelante.

## TCP/IP y el modelo OSI

Aún cuando los esquemas de las capas son diferentes, una comparación entre el conjunto de protocolos TCP/IP y el modelo de referencia OSI es válida debido a su estrecha relación.

El modelo OSI está diseñado para detectar y manipular errores en todos sus niveles. En cambio, TCP/IP realiza una detección de errores específicos en varias capas, efectuando en algunas un mínimo procedimiento de detección.

En el modelo OSI la mayor parte de la inteligencia de la red depende de ella misma, una red OSI es un sistema complejo e independiente; mientras que una red TCP/IP requiere de hosts inteligentes para participar en cada fase de esta (por lo tanto, una red Internet es una red con hosts inteligentes).

Existe una gran diferencia en la concepción de los esquemas: OSI fue y ha sido diseñado idealmente; en cambio TCP/IP es la historia de una evolución en donde los protocolos han sido creados y modificados conforme resolvían problemas.

Como puede verse en la figura 1.3b, TCP/IP presupone independencia del medio físico de comunicación, sin embargo existen estándares bien definidos a los niveles de *Liga de Datos* y *Físico* de OSI, que proveen mecanismos de acceso a los diferentes medios y que en el modelo TCP/IP deben considerarse.

<b>Aplicación</b>						
<b>Presentación</b>	TELNET	FTP	SNMP	SMTP	DNS	HTTP
<b>Sesión</b>						
<b>Transporte</b>	TCP					
<b>Red</b>	IP					
<b>Liga de Datos</b>	802.2					X.25
	802.3	802.5		LAPB		LLC/SHAP
<b>Física</b>	Ethernet	Token Ring	FDI	Línea Síncrona WAN	SONET	

Figura 1.3b. Modelo OSI y su relación con el Conjunto de Protocolos TCP/IP.

Para que la información fluya a través de estas capas, debe pasar por un proceso de encapsulamiento, es decir, los mensajes o información recibidos por la capa de TCP son reunidos con un encabezado en un paquete llamado "segmento de TCP", este segmento es entregado a la capa IP, en donde se le agrega un "encabezado de IP" y el paquete llamado "datagrama de IP" es creado.

Observe la siguiente figura:

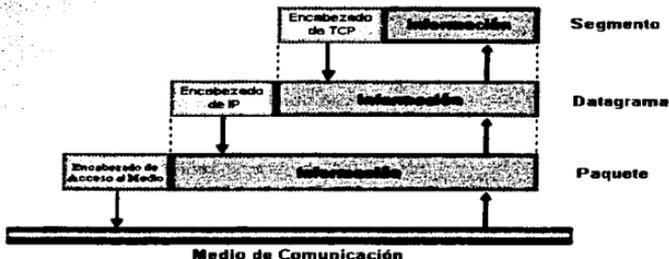


Figura 1.3c. Proceso de encapsulamiento de información.

El paso final incluye el encapsulamiento del datagrama de IP en paquetes de información creados para la capa de acceso al medio de comunicación, por la cual viajarán.

## Transmission Control Protocol y User Datagram Protocol

Los protocolos de la capa de transporte son: TCP (Transmission Control Protocol) y UDP (User Datagram Protocol), ambos utilizan el servicio de entrega de paquetes de IP, y pueden distinguir entre múltiples procesos en la misma máquina usando un número de puerto.

TCP provee un servicio "confiable" de entrega de paquetes de información "orientado-a-conexión", es decir, TCP se encarga de dar la ilusión de que la comunicación entre dos computadoras es de *punto-a-punto* con un flujo continuo de información, a diferencia de IP, donde la información fluye en paquetes y puede ser retransmitida varias veces antes de alcanzar su destino.

La conexión es muy importante ya que permite a un puerto local dar servicio a muchos puertos remotos concurrentemente, esta es la base del modelo de aplicación *cliente-servidor* que es usado y del cual se hablará más adelante.

La comunicación *punto-a-punto confiable* indica que TCP acepta la responsabilidad de la secuencia de datos, validación, y si es necesario, retransmisión; la aplicación o proceso que use los servicios de TCP, NO necesita preocuparse de todo lo anterior, puede asumir que los datos que envía serán recibidos íntegros, en el orden exacto en el que fueron enviados.

Otra de las responsabilidades de TCP es el control del flujo, el cual es un mecanismo que previene al transmisor de enviar datos más rápido de lo que el receptor pueda manejar.

UDP proporciona la capacidad de acceder a los puertos, a diferencia de TCP (con servicios *"sin-conexión"* y *"no-confiables"*). Muchas aplicaciones necesitan direccionar a IP y el acceso a puertos de TCP, pero manejando ellas mismas la verificación de los datos. También es usado por aplicaciones que solamente envían mensajes cortos y pueden enviar nuevamente los mensajes si la respuesta no llega en corto tiempo.

TCP toma la información que se quiere transmitir, la divide en segmentos y numera cada uno de estos, de tal manera que el receptor pueda verificar la llegada de los mismos y colocarlos en el orden correcto.

## **Internet Protocol**

IP (Internet Protocol) es llamado la base tecnológica del conjunto de protocolos TCP/IP. Las funciones que realiza son las siguientes:

Provee un servicio de entrega de datagramas *"sin-conexión"*, llamado así porque NO se lleva a cabo una coordinación entre el punto transmisor y el punto receptor. Cada paquete es tratado independientemente, los cuales pueden llegar en desorden y hasta podrían NO llegar.

La entrega *"sin-conexión"* análogicamente corresponde a poner una carta en el buzón: se deposita (datagrama), se olvida uno de ella y se asume que el servicio postal (Red IP), entregará la carta (datagrama) a su destino.

Este servicio *"sin-conexión"* es *"no-confiable"* ya que IP NO puede garantizar la entrega, sin embargo es llevado a cabo con el mejor esfuerzo, es decir, los datagramas no son descartados fácilmente, estos pueden no ser entregados por falta de recursos o por una falla en el hardware de la red.

### **1.4 Direccionamiento**

El direccionamiento de IP determina rápidamente si una dirección IP dada por la capa de transporte pertenece a la red local o a otra red. Las direcciones IP son números de 32 bits divididos en 4 octetos. Cada dirección es la combinación del identificador único de la red y el identificador único de la máquina. IP es responsable de la selección de la ruta por el que viajan los datos, esto es comúnmente llamado *"enrutamiento"*.

IP también incluye un conjunto de reglas que define cómo se procesarán los paquetes de información, incluyendo cuándo generar mensajes de error y cuándo se descartan datagramas. Parte de este proceso incluye la "Fragmentación de Datos" y el "Reensamble", aunque IP realiza esta función sólo cuando el hardware lo requiere.

Para emplear Internet, una de las primeras cosas que se deben saber es qué es una dirección, ya que para acceder a los diversos servicios de la red siempre habrá que especificar al menos una.

Todas las computadoras conectadas a Internet tienen una dirección numérica única e irreplicable, llamada *dirección IP*, la cual sirve para poder comunicar a unas máquinas con otras, la dirección no se asigna arbitrariamente, se debe hacer una petición al Network Information Center (NIC), organismo responsable de la administración de las direcciones de toda la red Internet; generalmente uno como usuario no se tiene que preocupar por esto, el administrador de la red es quien tiene que hacerlo.

La dirección IP se compone de cuatro partes: *país*, *organización*, *subred*, y *número de la computadora*. La secuencia de los datos es jerárquica, o sea que el primer dato es el más general y el último, el más particular. Primero es el país donde se localiza la computadora (conocido también como *dominio*), seguido de la organización o empresa que tiene la máquina, después sigue la subred donde se localiza dentro de la organización, y por último el número que se le asignó a la máquina. Hay que señalar que NO todas las máquinas siguen esta convención.

A cada uno de estos datos le corresponde un número, y juntos forman la dirección completa al escribirse separados por un punto.

Por ejemplo la dirección: **132.248.120.27** donde:

<b>132</b>	es el número asignado a México.
<b>248</b>	es el número asignado a la UNAM.
<b>120</b>	es el número asignado a la subred de la DGSCA.
<b>27</b>	es el número asignado a la máquina dentro de la red de la Dirección de Cómputo para la Docencia.

Las direcciones numéricas son difíciles de recordar, así que se desarrolló una forma basada en nombres. Al principio para darle nombre a una computadora había un archivo llamado HOSTS.TXT, el cual contenía una tabla de equivalencias entre nombres y números. Por ejemplo, para la máquina del ejemplo anterior, su entrada en el archivo sería:

>S **132.248.120.27 morsa.dgscsa.unam.mx**

donde:

<b>morsa</b>	es el nombre que se le ocurrió a mi jefe, responsable del departamento (completamente arbitrario).
<b>dgscsa</b>	nombre de la dependencia.
<b>unam</b>	en la UNAM.
<b>mx</b>	en México.

El nombre va en orden inverso a la dirección numérica, sin embargo denota a la misma máquina. Como era imposible mantener un archivo que contuviera todos los nombres de las computadoras del mundo, se creó el sistema de dominio de nombres o **DNS** (Domain Name Services); este sistema es una base de datos distribuida en todo el mundo que contiene equivalencias entre nombres y números, y con el se elimina el uso del archivo HOSTS.TXT y en su lugar se utilizan los servidores de nombres (computadoras dedicadas), en donde se efectúa una búsqueda de la equivalencia entre el nombre y una dirección, y en caso de no encontrarla, se efectúa una petición a otros servidores.

No todos los nodos de Internet tienen la misma estructura en la dirección IP, sobre todo los que se encuentran en EUA, estas computadoras poseen un código propio, sin embargo han dejado de emplearse, también son jerárquicas, pero no se tiene que especificar el país (dominio), en su lugar se escribe el tipo de organización que tiene el nodo. Por ejemplo:

*tron.sis.uiowa.edu*

La cual es una dirección IP de una computadora llamada tron, dentro del departamento de sistemas en la Universidad de Iowa, la cual es una institución educativa.

Los dominios "organizacionales" más comunes en EUA son:

<b>.edu</b>	Para universidades e Instituciones educativas. ejemplo: <i>nyu.edu</i> (New York University).
<b>.com</b>	Para empresas comerciales que dan servicios dentro de Internet. ejemplo: <i>ftp.cdrom.com</i>
<b>.gov</b>	Para instituciones gubernamentales. ejemplo: <i>nasa.gov</i>
<b>.org</b>	Para organizaciones privadas no lucrativas. ejemplo: <i>eff.org</i> (Electronic Frontier Foundation).
<b>.net</b>	Para redes de enlace dentro de Internet. ejemplo: <i>nsf.net</i> (National Science Foundation).
<b>.mil</b>	Para Instituciones militares.

A medida que Internet se convirtió en una red de alcance mundial, fue necesario dar a cada país la responsabilidad de sus propios nombres. Por esta razón, existe un conjunto de dominios de dos letras correspondientes a los dominios "geográficos" para países.

Por ejemplo:

<b>.mx</b>	México.
<b>.au</b>	Australia.
<b>.ca</b>	Canadá.
<b>.fr</b>	Francia.
<b>.uk</b>	Reino Unido.

Estos además pueden tener sub dominios como: *ac.uk* para instituciones académicas o *co.uk* para instituciones comerciales.

## 1.5 Modelo Cliente-Servidor

La mayoría de las aplicaciones TCP/IP están creadas bajo el modelo Cliente-Servidor, el cual puede verse como la extensión lógica de la programación modular, la cual se fundamenta en separar un gran programa de aplicación en sus partes constitutivas o módulos, de esta manera existe la posibilidad de un análisis y desarrollo más sencillo, detallado y dinámico, así como un mejor mantenimiento. El módulo que hace peticiones o requerimientos de un servicio, se denomina "cliente", y al que se le hacen peticiones, se llama "servidor" es decir, el que provee el servicio.

El programa servidor es referido como "pasivo", debido a que generalmente espera que le sea requerido un servicio. La mayoría de las aplicaciones en un servidor son implementadas en programas o rutinas llamadas *daemons*, que son llamados cuando el sistema de computo es inicializado (o booted), y se ejecutan en modo *background*, es decir en espera de una petición.

El programa cliente es "activo", es decir que inicia su actividad al realizar una petición al servidor. La interfaz de usuario para obtener algún servicio es usualmente un programa cliente (por ejemplo ftp o telnet).

Existen varias razones para la existencia de este modelo:

- Se tienen clientes y servidores corriendo con hardware y software apropiados para sus funciones respectivas; por ejemplo, un sistema servidor de bases de datos, corre en una plataforma especialmente diseñada y configurada para realizar consultas, o los servidores de archivos corren en plataformas que poseen elementos especiales para el manejo de archivos.
- Permite una economía de recursos, ya que no es necesario tener varios equipos potentes, se puede sumar la especialización de tareas de varios equipos pequeños, para dar un mejor rendimiento; por ejemplo, en vez de tener un sólo servidor que sea: servidor de archivos, de impresión, de correo, de servicios de nombres; podemos tener varios de menor capacidad dedicados a tareas específicas.
- El tráfico en la red se distribuye y puede disminuir, aumentando su rendimiento.
- Permite tener un ambiente heterogéneo, ya que por ejemplo, el hardware, así como el sistema operativo del cliente y del servidor generalmente NO son el mismo.
- Permite una alta escalabilidad, que puede ser horizontal o vertical. La escalabilidad horizontal, significa agregar o quitar clientes con muy poco impacto al rendimiento del sistema. Escalamiento vertical es migrar a servidores más grandes y potentes.

*Específicamente, el cliente es un programa que envía mensajes o peticiones al proceso que se ejecuta en el servidor, pidiéndole que realice una tarea (o servicio), el cliente usualmente maneja una parte de la aplicación, es decir que la interfaz de usuario empleada se encarga de validar datos, enviando las peticiones al servidor; así, el cliente resulta ser el *front-end* con el que el usuario interactúa.*

El servidor es un proceso, que responde a las peticiones del cliente, realizando las tareas requeridas, este recibe las peticiones del cliente, ejecuta búsquedas y actualizaciones a bases de datos, administra la integridad y respuesta de las peticiones del cliente. El proceso servidor puede correr en otra máquina en la red. También actúa como el controlador de los recursos compartidos en la red, tales como bases de datos, impresoras, discos, enlaces de comunicación, procesador, etc. Se trata del *back-end* del sistema.

Algunos ejemplos comunes de éste tipo de modelo, son sistemas manejadores de bases de datos, donde el cliente envía peticiones SQL (Structured Query Language) y el servidor regresa el resultado de estas. NFS es otro ejemplo, ya que un cliente efectúa una petición a una computadora para compartir recursos de disco, es entonces que el servidor del sistema retorna el acceso a los sistemas de archivos. Otro ejemplo es un sistema X-Windows, donde el cliente tiene la capacidad de ejecutar programas que se muestran en el servidor.

## 1.6 Aplicaciones y Servicios

Existe una gran cantidad de programas de aplicación que proporcionan numerosos servicios en Internet, los cuales emplean protocolos de transporte para comunicar diversos hosts en toda la red; algunos de estos servicios, ubicados en la capa de aplicación son:

### Telnet

Es un protocolo de terminal remota (estándar de los protocolos TCP/IP). Este protocolo permite tener acceso remoto a otro host, por medio de una conexión TCP. El hecho de que un usuario este introduciendo comandos desde la red y no desde una terminal conectada directamente, es transparente al sistema operativo del host remoto.

El cliente es el programa que el usuario ejecuta (por medio del comando *telnet*). El servidor es un programa *daemon* (usualmente denominado *telnetd*) en el host remoto. El protocolo Telnet realiza tres servicios básicos:

1. Define la interfaz Network Virtual Terminal (NVT), para sobreponerse a las dependencias del sistema. NVT permite al cliente y al servidor implementados sobre diferentes sistemas operativos comunicarse definiendo una interfaz común.
2. Provee un proceso para que el cliente y el servidor puedan negociar sobre varias opciones (por ejemplo, el conjunto de caracteres a usar).
3. Trata ambos puntos de la conexión simétricamente. Esto significa que el cliente NO debe estar necesariamente conectado a una terminal físicamente (esto permite hacer Telnet a otra máquina, y desde ahí otro Telnet, etc.).

## File Transfer Protocol (FTP)

FTP es un estándar del conjunto de protocolos TCP/IP que efectúa la transferencia de archivos entre host a través de la red. FTP es implementado como una aplicación cliente-servidor, donde el usuario ejecuta el programa cliente (por medio del comando ftp), y el programa servidor es un *daemon* (*ftpd*). FTP provee servicios como:

- Interfaz de usuario interactiva.
- Especificación del formato. El usuario puede seleccionar el tipo de transferencia de datos a emplear (binario o texto), y el formato del texto: ASCII o EBCDIC.
- Autenticidad. El usuario debe conectarse por medio de una cuenta disponible en el host remoto y validar la cuenta por medios de una palabra clave (password).

FTP es independiente del sistema operativo, a diferencia de rcp o rlogin (servicios NO estándares del conjunto de protocolos TCP/IP), basados en UNIX.

## Simple Mail Transfer Protocol (SMTP)

El correo electrónico es uno de los servicios de red más ampliamente usados. Existen básicamente dos estándares para proveer servicio de correo electrónico sobre TCP/IP en Internet:

1. Formato Mail message, especificado por el RFC 822
2. Mail Exchange, especificado por el Simple Mail Transfer Protocol, en el RFC 821.

El primero, consiste de un encabezado y un cuerpo, donde el encabezado posee palabras claves y valores separados por punto y coma. El formato del cuerpo es determinado por el emisor, dando flexibilidad al correo electrónico.

SMTP es una aplicación implementada como cliente-servidor. Típicamente el cliente establece una conexión TCP con el servidor y espera. El servidor envía el mensaje READY FOR MAIL. Cuando el cliente recibe este mensaje, envía el mensaje HELLO y se identifica a sí mismo. Después de esto, el cliente envía el mensaje. Cuando la transferencia del mensaje ha sido terminada, se intercambian los papeles con el servidor, esto permite al servidor convertirse en cliente, tomar el control de la conexión y transferir mensajes. En todos los casos el receptor debe reconocer los mensajes, y el emisor NO debe borrar su copy del mensaje hasta que no haya recibido el reconocimiento del mensaje.

Otras aplicaciones comunes de TCP/IP son: el Servicio de Nombre de Dominios (DNS), Protocolo de Ruteo de Información (RIP) y el NFS (Sistemas de Archivos de Red).

Además de estas aplicaciones existen otras, por ejemplo NIS (Network Information Service), conocido anteriormente como *Yellow Pages*, de manera similar al DNS, es empleado para distribuir información.

Gran parte del poder de Internet radica en los servicios que posee. Muchos usuarios desean, más que conocer con que máquinas pueden conectarse, saber que usuarios existen y los proyectos que se desarrollan en cada lugar, para ello se han implementado directorios de nombres, el primer servicio de esta naturaleza fue *Whois*, el cual contaba con una base de datos centralizada que permitía consultas remotas, dando a conocer que usuarios estaban registrados y en que computadoras. Además existe otro servicio llamado *X.500*, el cual es un manejador de directorios de nombres, y forma parte de la implementación del modelo OSI.

Se han generado nuevas herramientas dentro de Internet, y se puede afirmar que estas forman un nivel adicional en cuanto a la forma de mostrar la información en la "red de redes", por ejemplo: *Archie*, el cual es un servicio prestado originalmente por la Universidad de McGill en Montreal, que tiene por objeto localizar información de todos los servidores FTP *anónimo*. *Gopher* era otro servicio muy popular, desarrollado originalmente por la Universidad de Minnesota, para ayudar a encontrar información en su campus.

Este servicio permite acceder a información, está diseñado para trabajar en forma distribuida y basado en el modelo *cliente-servidor*, donde:

- Se puede organizar la información para atender las necesidades locales de los usuarios de un gran número de entidades.
- La mayor parte de los servicios de Internet son compatibles con Gopher, lo que amplía el horizonte de información y de servicios disponibles para el usuario.
- La información puede residir en servidores diferentes y completamente autónomos.
- El enlace de un servidor a otro es completamente transparente al usuario.

Estas características dan como resultado:

- Una comunicación confiable y absoluta entre clientes y servidores.
- Interacciones cooperativas cliente-servidor, originadas por un cliente.
- Distribución del procesamiento de la aplicación entre un cliente y su servidor.
- Control implantado en el servidor sobre servicios y los datos que un cliente puede solicitar.
- Criterios de decisión instrumentados en el servicio, sobre situaciones de conflicto en requisiciones de clientes.

Otro servicio es el denominado *WAIS*, (*Wide Area Information Service*), servicio de información de área extendida. *WAIS* es un protocolo diseñado para acceder a información indizada, bajo un formato especial que permite una recuperación de texto completo (esta constituido por un indizador, un servidor y un cliente).

Existen muchas maneras de indizar algún tipo de dato (como es el caso en los índices en **Archie** o **Whois**). Al igual que Gopher, WAIS es un sistema distribuido de búsqueda de información cuya arquitectura se basa en el modelo cliente-servidor y en el estándar Z39.50 (protocolo de computadora a computadora, estándar de ANSI), para peticiones bibliográficas de información, este permite encontrar y acceder a recursos en la red.

*Sin lugar a duda una de las aplicaciones más importantes y que ha originado la inquietud de escribir este trabajo es: **World Wide Web** del cual se hablará extensamente en los siguientes capítulos, ya que además de ser el detonador de la enorme difusión a nivel mundial de **Internet**, es el medio de comunicación que está transformando a la sociedad, con un impacto mayor al de la televisión o el cine, aunado al enorme crecimiento que presenta y a los avances tecnológicos que se esperan en pocos meses, se puede afirmar con certeza que **WWW transformará por completo la forma de vida de la humanidad**.*



## Capítulo II

# World Wide Web

---

Gracias a los avances tecnológicos en materia de telecomunicaciones como son: enlaces satelitales, fibra óptica, computación remota y móvil, telefonía celular, etc., se ha permitido la conectividad mundial.

**WWW** o simplemente **Web**, es un medio funcional para que la gente de todo el mundo localice información y comparta en forma general el conocimiento humano, gracias a la facilidad de su uso comparado con otros sistemas de recuperación de información como: *Archie*, *Gopher* o *WAIS*, este medio de comunicación enlaza a la mayoría de los recursos, servicios y aplicaciones que existen en Internet. El éxito del Web radica en la integración que presenta de variados formatos de información como: imágenes, texto, audio o video. Al emplear el Web, sin esfuerzo alguno se logra visualizar información de servidores localizados físicamente en cualquier parte del mundo, aplicaciones de sistemas con *FTP* y *Telnet*, y formatos de información como archivos y documentos; además, las cualidades multimedia llaman la atención de los usuarios en general.

World Wide Web es un sistema de intercambio de información multimedia desarrollado en el CERN (Centro Europeo de Investigación Nuclear) de Ginebra en 1989. Los resultados iniciales obtenidos por el físico **Tim Berners-Lee** durante el intercambio de documentación interna de proyectos, fue tan exitoso que a los pocos meses de su creación, se comenzó a difundir con rapidez a otras entidades, tanto locales como internacionales. En 1993, existían alrededor de 200 *servidores Web*; hoy Internet posee más de 80 mil distribuidos alrededor del mundo.

Las primeras difusiones de este novedoso sistema de intercambio de información, dieron lugar a la creación del **WWW Consortium (W3C)**, organización que agrupa a instituciones interesadas en el desarrollo de este medio de comunicación. El consorcio propone nuevas funciones del sistema, evalúa propuestas e implementaciones de herramientas, entre otras muchas actividades.

**WWW** son las iniciales de **World Wide Web**, lo que se puede traducir al español como simplemente: *"red mundial amplia"* ó *"red global"*; y éste es el propósito expresado por W3C: *"hacer de WWW un sistema multimedia de recuperación de información a nivel global"*, y para el CERN (creadores del Web): *"World Wide Web es el universo de información accesible por la red, todo un mundo de conocimiento humano. Es una*

*iniciativa que empezó en el CERN, y ahora con muchos participantes. Posee un cuerpo de software, y un conjunto de protocolos y convenciones. WWW usa hipertexto y técnicas multimedia para hacer a Internet fácil de recorrer, buscar y contribuir"*

El Web se compone de diversos elementos donde interactúan: **Servidores**, los cuales contienen sitios con información que puede ser distribuida al mundo; **Clientes**, a través de los cuales los usuarios pueden navegar a lo largo y ancho de la "red de redes", empleando para ello **visualizadores** (aplicación-cliente), populares como: Navigator, Explorer o tal vez los antiguos Mosaic y Lynx (visualizador modo texto). También se compone de **Agentes**, es decir, programas o scripts empleados para facilitar la comunicación y proveer control de acceso.

En el CERN se desarrolló el primer *servidor Web* y un primer *programa cliente* (llamado también *visualizador*), pero el segundo gran impacto en el desarrollo del Web fue la aparición de **Mosaic**, primer visualizador gráfico, desarrollado por **NCSA** (National Center for Supercomputing Application) en EUA y que dio la dimensión multimedia a WWW. La importancia de Mosaic era hasta hace algunos años tal, que con frecuencia se identificaba con este nombre al Web y hacia referencia a Internet.

El Web emplea el protocolo **HTTP** (HyperText Transfer Protocol), y un sistema de descripción de archivos incorporado en el lenguaje de marcación de hipertexto (**HTML**), este lenguaje será abordado profundamente en los dos siguientes capítulos.

## **2.1. HTTP**

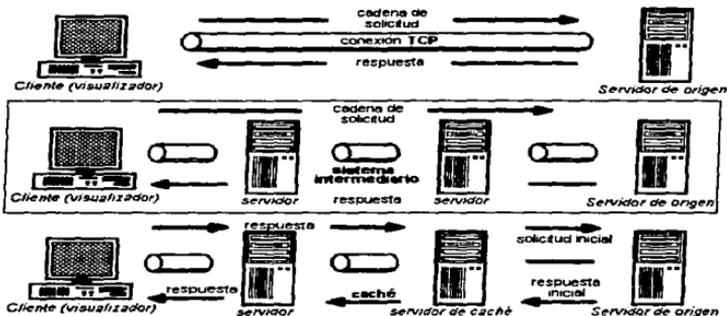
HTTP es el protocolo de transferencia de hipertexto que emplea WWW, el nombre es un tanto engañoso, HTTP NO es un protocolo para transferir hipertexto, se trata de un *protocolo para transmitir información con la eficiencia necesaria para hacer enlaces hipertexto*. Los datos transferidos por el protocolo pueden ser texto sencillo (texto plano), hipertexto, audio, imágenes, video, o cualquier información accesible a través de los diversos servicios Internet.

*HTTP es un protocolo cliente-servidor orientado a transacciones*. Para asegurar la confiabilidad emplea TCP: se trata de un protocolo "sin estado", es decir, trata cada transmisión en forma independiente. Una implementación típica puede crear una nueva conexión TCP entre el cliente y el servidor para cada transmisión de información (transacción, en este caso), luego entonces, se puede terminar la conexión tan pronto como se complete la transacción, sin embargo, la especificación NO requiere esta relación de uno a uno entre el tiempo de vida de la transacción y la conexión; por ejemplo, la conexión puede permanecer abierta de manera que se puedan hacer más transmisiones.

La naturaleza "sin estado" de HTTP se encuentra bien definida a su aplicación típica: *una sesión Web normal, involucra recuperar una secuencia de páginas y documentos, esta*

secuencia es tratada rápidamente (idealmente), y las ubicaciones de varias páginas y documentos pueden ser distribuidas ampliamente entre un número de servidores localizados alrededor del mundo.

La figura 2.1a muestra tres ejemplos de operaciones HTTP. El cliente (o usuario), es un visualizador Web que efectúa peticiones, el servidor origen es el servidor en el que reside un recurso, al cual se quiere acceder, por ejemplo: un servidor donde se encuentre localizado un Sitio Web o Web Site. El caso más sencillo es en el cual un cliente establece una conexión directa con un servidor origen, el cliente abre una conexión TCP (la cual permanece de principio a fin) entre el cliente y el servidor "rápidamente", el cliente entonces emite una solicitud HTTP, esta solicitud consiste en un comando específico referido, como un *método* (ver más adelante), un URL y un mensaje que contiene los parámetros de la solicitud, información sobre el cliente, y posiblemente información adicional del contenido.



HTTP soporta operaciones a través de conexión directa, sistemas intermediarios o caché.

Figura 2.1a. Tipos de transferencia HTTP.

Cuando el servidor recibe la solicitud intenta realizar la acción pedida, devuelve una respuesta HTTP, esta incluye información de estado, un código de éxito/error y un mensaje que contiene información sobre el servidor, e información sobre la respuesta en sí, así como el posible contenido del cuerpo, es entonces que la conexión TCP se cierra.

La parte media de la figura 2.1a, muestra un caso en el cual no existe una conexión TCP de principio a fin entre el cliente y el servidor origen, en lugar de esto, existen uno o más sistemas intermediarios con conexiones TCP entre sistemas adyacentes lógicos. Cada sistema intermediario actúa como un transmisor o retransmisor, de manera que una

solicitud iniciada por el cliente es enviada a través de sistemas intermediarios hacia el servidor, y la respuesta de este es transmitida de regreso al cliente.

La especificación HTTP define tres formas de sistemas intermediarios:

- Sustituto
- Portal
- Túnel

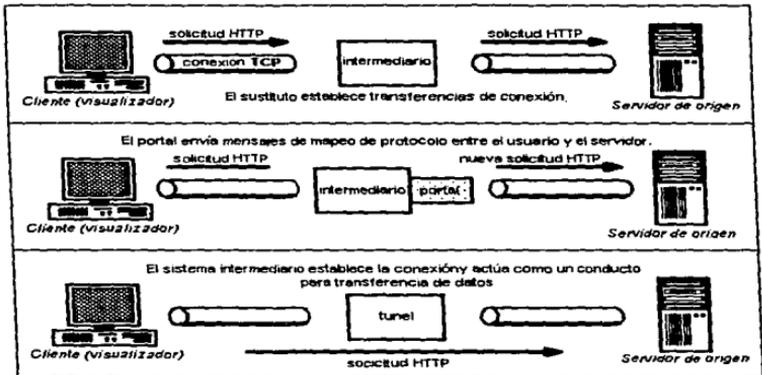


Figura 2.1b. Sistemas HTTP Intermediarios.

Un sustituto actúa en nombre de otros clientes y presenta solicitudes de estos a un servidor, existen varios casos que requieren el uso de un sustituto, uno de ellos es cuando el sustituto actúa como un intermediario a través de un *firewall*. En este caso, el servidor debe autenticarse a sí mismo ante el usuario, es entonces que el sustituto acepta las respuestas, una vez que estas han sido validadas por del firewall (observe la figura 2.1b).

Otro caso que se presenta es al manejar diferentes versiones de HTTP<sup>1</sup>, es decir si el cliente y el servidor se encuentran ejecutando diferentes versiones del protocolo, es entonces que el sustituto puede implementar ambas versiones y realizar el mapeo requerido para lograr establecer el canal de comunicación.

<sup>1</sup> La información más reciente enviada al Grupo de Estándares de Internet (como estándar propuesto), es el borrador 05 de la versión HTTP 1.1 (junio 1996).

Un portal es un servidor que se presenta (ante un cliente), como si se tratara de un servidor origen y actúa en nombre de otros servidores que podrían NO ser capaces de comunicarse directamente con el cliente. Al igual que con un sustituto, un portal maneja transferencias a través de un firewall, en este caso, el cliente debe autenticarse a sí mismo ante el sustituto, el cual a su vez puede pasar la solicitud al servidor.

Otro caso común es trabajar con un servidor que NO sea HTTP. La gran mayoría de los visualizadores (como Navigator de Netscape), tienen integrada la capacidad de establecer comunicación con servidores que emplean protocolos diferentes a HTTP, tales como servidores FTP, Gopher, mail, etc. Esta capacidad de manipular múltiples protocolos, también puede ser proporcionada por un portal. Un túnel es simplemente un punto de transmisión entre dos conexiones TCP, los mensajes HTTP pasan sin cambios, sólo en el caso de que exista una sola conexión HTTP entre el usuario-cliente y el servidor de origen. Los túneles son comúnmente empleados cuando existe un sistema intermediario entre el cliente y el servidor.

Otro tipo de operación de HTTP es que interactúa con el caché. El caché permite almacenar solicitudes y respuestas, con la finalidad de manejar nuevas solicitudes. Si llega una nueva petición que emplea la misma almacenada, entonces el caché puede proporcionar la respuesta almacenada en vez de tener acceso al recurso indicado en el URL, evitando así la congestión de tráfico en la red. El caché puede operar en un cliente, en un servidor o en un sistema intermediario diferente al túnel. En la figura 2.1a, un servidor a efectuado una petición, la cual es respondida por el caché de una transacción de *solicitud/respuesta* efectuada, de manera que una nueva solicitud correspondiente al cliente NO necesita viajar por toda la cadena de hosts establecida por TCP hasta el servidor origen; en vez de eso se hace uso del caché; NO todas las transacciones pueden hacer uso de este, y un cliente o un servidor por su configuración, puede determinar si cierta transacción debe emplear el caché durante una cantidad determinada de tiempo.

## Mensajes HTTP

Los mensajes HTTP comprenden dos tipos: *solicitud* y *respuesta*. Un mensaje de solicitud se envía por un cliente a un servidor para iniciar alguna acción, entonces se devuelve un mensaje de respuesta por un servidor a un cliente, en respuesta a la solicitud. Algunas acciones o *métodos* posibles son:

- **GET:** solicitud para recuperar información.
- **POST:** solicitud para aceptar la entidad anexa como un nuevo subordinado al URL identificado.
- **PUT:** solicitud para aceptar la entidad anexa y almacenarla bajo el URL proporcionado (este es un nuevo recursos con un nuevo URL), o inclusive puede ser un reemplazo del contenido de un recurso ya existente con un URL.
- **DELETE:** solicita que el servidor de origen borre un recurso.

Una respuesta incluye un cuerpo de mensaje que contiene información basada en hipertexto, adicionalmente el mensaje de respuesta debe especificar un **código de estado**, el cual indica la acción tomada en la solicitud correspondiente. Los **códigos de estado** se encuentran organizados en las siguientes categorías:

- **Informational:** La solicitud se ha recibido y el procesamiento continua (ningún cuerpo de entidad acompaña esta respuesta).
- **Successful:** La solicitud se ha recibido, comprendido y aceptado con éxito.
- **Client error:** Las solicitudes contienen un error de sintaxis o la solicitud no puede completarse.
- **Server error:** El servidor no pudo completar una solicitud aparentemente válida.

Una vez establecida la conexión, HTTP se encarga de determinar el objetivo de la misma (servir o recibir documentos HTML). Existen cuatro **fases secuenciales** en una transacción HTTP: *conexión, petición, respuesta y desconexión*.

- **Connection:** El cliente intenta conectarse a la *dirección Web* ó *URL* del servidor (en realidad al software servidor que se está ejecutando en la máquina remota). El software cliente reconoce el momento en que la conexión ya se ha realizado, pasando a la siguiente fase.
- **Request:** Una vez completada la conexión, el cliente envía una petición al software servidor para ejecutar una acción específica. Esta petición, incluye la siguiente información: **Request method**, que se refiere a la manera en que se comunicarán cliente y servidor, así como la forma en que ambos llegarán a intercambiar información; **Request URI**, refiriéndose al recurso exacto al que se aplicará la petición del cliente; **Request header files**, que contiene información adicional necesaria que determina el comportamiento final del servidor.
- **Response:** La acción real de una conexión HTTP es realizada en esta fase; aquí, el software del servidor intenta satisfacer la petición del cliente. Se responde al cliente con una línea de estado, incluyendo información acerca del servidor y los datos solicitados, o bien, un mensaje de error.
- **Close:** En esta fase, la conexión entre el cliente y el servidor es terminada. Por lo general, cada transacción comienza con el cliente que establece una conexión, y termina con el servidor, cerrando la misma después de atender la petición del cliente.

## 2.2. Componentes de WWW

WWW basa su diseño y funcionamiento en los siguientes puntos:

- La idea de un mundo sin límites para la información se basa en que todo elemento o recurso disponible, se encuentre identificado por una clave que permita acceder a este y recuperarlo.
- Un sistema de direcciones (**URI/URL**) hace posible este mundo, al margen del conjunto de protocolos TCP/IP empleados.
- Un protocolo de red (**HTTP**) empleado por los servidores WWW proporciona funciones y recursos que no brinda algún otro.
- Un lenguaje (**HTML**) que reconocen universalmente todos los clientes WWW y que se usa para transmitir información a través de la red.
- El cuerpo de datos e información disponible en Internet emplea todos o alguno de los elementos antes mencionados.

Los **URI** (Universal Resource Identifier) o **URL** (Universal Resource Locator) son cadenas que identifican objetos en el Web, a través del conjunto de protocolos TCP/IP. Estos son "*universales*", lo cual quiere decir que se pueden emplear para identificar objetos casi de cualquier naturaleza y exportarlos con cualquier protocolo. De ahí que aunque se asocie su uso a WWW no deba restringirse a éste y sean capaces de direccionar igualmente otros componentes en cualquier protocolo como FTP, Gopher, SMTP, etc.

**HTTP** fija las condiciones necesarias que permiten hacer uso de la forma de presentación de información en WWW, trabajando de igual manera con texto, gráficos o sonidos. HTTP es muy parecido al protocolo FTP de transferencia de archivos y NNTP (Network News Transfer Protocol) usado por News. Los cuales trabajan sobre la capa de aplicación del conjunto de protocolos TCP/IP.

La conexión HTTP, al contrario de lo común en otros protocolos se mantiene únicamente durante el tiempo necesario para efectuar una operación. La petición desde un *cliente* comienza indicando un código de operación o método y una *dirección URL* (los principales métodos son GET, PUT, POST, etc., vistos anteriormente). Cuando se transfiere un *documento WWW* a través de la red, también se envía metainformación sobre este, muchas veces esta consiste en identificadores **MIME** (Multipurpose Internet Mail Extensions) que definen los tipos de componentes.

**HTML** (HyperText Markup Language) es un *sistema unificador de información* que permite integrar en un mismo documento objetos de distintas naturalezas. Deriva del estándar **SGML** (Standard Generalized Markup Language) o Lenguaje Normal Generalizado de Marcapción (norma ISO), derivada de un estándar antiguo (GML).

SGML permite que la estructura de un documento pueda ser definida en base a la relación lógica de sus partes. Esta estructura puede ser validada por una Definición de Tipo de Documento (DTD - Document Type Definition). La *norma SGML* define la sintaxis del documento y la semántica que definen los distintos niveles de HTML<sup>2</sup>. El lenguaje muestra que se originó en sistemas donde el texto era el contenido principal y el marcado era la excepción; puede ser usado en el futuro para otros propósitos como "*correo multimedia*"; también puede ser empleado como lenguaje en el que se podrán implementar secciones *multimedia* e *hipertexto* dentro de archivos escritos en otros lenguajes.

La *unidad de trabajo* de WWW es el **documento** (o *página*), normalmente escrito en HTML. Cada uno de estos documentos contiene texto con formato e indicaciones sobre el aspecto (tipo de letra, tamaño, etc.) y la disposición del texto (párrafos, listas, tablas, frames, etc.). Una mínima estructura organiza el documento en secciones más o menos fijas; también contiene imágenes, soportadas en varios formatos, que se pueden visualizar directamente en el documento o externamente por medio de un software adecuado. Este mismo sistema se emplea para objetos de todo tipo (imágenes, sonidos, etc.). Un *documento WWW* contiene referencias a otros documentos por medio del URL.

Esta última característica es la que ha hecho tan atractivo a este sistema de intercambio de información. El usuario común de Internet ve un documento con la apariencia de una página compuesta por letra proporcional, texto con formato, color, texturas, imágenes, gráficos y enlaces hipertexto. Con el simple uso del "*ratón*" (mouse), se muestra un nuevo documento que puede estar contenido en un servidor en el otro extremo del mundo. Esto es lo que se ha dado en llamar "*navegar en el Web*".

Para el usuario de Internet es muy fácil obtener un documento (obviamente conociendo el URL adecuado) y, desde allí, localizar información relacionada con el tema de su preferencia. De un punto a otro de Internet, se puede pasar con rapidez a temas interrelacionados pero que pueden acabar teniendo muy poco en común con el punto de partida, *este es el efecto típico de las aplicaciones hipertexto*, pero maximizado en una red de ámbito y fines globales.

## **2.3. Hipertexto e Hipermédia**

El hipertexto es un subconjunto de hipermédia, se refiere específicamente a documentos en donde los lectores pueden moverse de un lugar a otro, de una página a otra o entre documentos, desplazándose en forma aleatoria, *de manera NO secuencial ni lineal*, es decir, NO se accede a la información en forma tradicional, de principio a fin; las palabras, frases e imágenes del documento se convierten en enlaces a través del hipertexto que permiten viajar a una nueva posición en el mismo documento o en alguno otro, como

---

<sup>2</sup> Actualmente existen cuatro niveles: 0, 1, 2 y 3 (ver capítulo III, HTML).

resultado se pueden efectuar "enlaces" entre documentos y fuentes de diferentes medios de información localizados en miles servidores Web alrededor del mundo.

El hipertexto, tiene varias ventajas sobre el texto normal:

- Facilita la navegación en documentos muy largos. En ocasiones los volúmenes de manuales de ciertos productos, son tan grandes para ser impresos, que se prefiere entregar éstos en un formato hipermedia, así es mucho más fácil que el lector pulse con el mouse una palabra determinada y se mueva inmediatamente a ese tema en el documento, en vez de localizar el volumen al que se hace referencia y encontrar la página adecuada en la versión impresa.
- Además de la velocidad de uso, el hipertexto ayuda a los lectores a explorar nuevas ideas y localizar nuevas fuentes de información a medida que se desplaza de un lugar a otro.
- Los lectores se convierten en exploradores y toman decisiones de "navegación" acerca de los temas que desean investigar. El hipertexto permite al usuario común, decidir de manera precisa qué información es la más importante en un momento determinado.

En hipermedia, los enlaces son conexiones visuales a gráficas o fotografías, mensajes de audio o video, así como a texto normal e hipertexto. *En forma general hipertexto e hipermedia, son herramientas que posibilitan una fácil maniobra, de los inmensos y casi inagotables recursos disponibles en los servidores Web de Internet.*

## 2.4. Servidores y Clientes en WWW

Los recursos de Internet y WWW son muy grandes. Millones de documentos se encuentran disponibles alrededor del mundo y más de *80 mil servidores Web* soportan información interactiva. La comunicación a través del Web sigue el modo de operación propio del *modelo cliente-servidor*; en un extremo de la comunicación tenemos una computadora sobre la que corre la *aplicación* en el servidor HTTP (servidor Web), y en el otro existe con un *programa cliente*, comúnmente llamado "visualizador" o inclusive "browser".

Cuando se quiere mostrar o ver una página, el cliente envía al servidor una petición (cuya dirección conoce a través del URL del documento), el servidor recupera el documento (en el servidor HTTP) y lo envía al cliente; si éste no existiera, o el cliente no estuviera autorizado a leerlo, o si se produce cualquier error en la transferencia, el servidor aborta la operación y devuelve un código y mensaje de error. Este documento puede contener referencias a imágenes y otros objetos, el cliente debe ir resolviendo estas, a la vez que las va decodificando, invoca a los servidores apropiados (si se presenta el caso), que

pueden ser HTTP o algún otro, dependiendo del protocolo indicado en el URL de cada objeto.

Los servidores son programas comerciales que operan de acuerdo al protocolo HTTP. Un servidor opera en una máquina normalmente con un sistema operativo robusto como UNIX (existen versiones VMS, Solaris, etc.), Mac, Windows, Linux, entre otros. La computadora en la cual se encuentra el servidor Web administra el conjunto de documentos que se quieren difundir.

Los servidores permiten establecer condiciones suficientes para restringir el acceso a la información contenida en estos, a través de diversos programas y mecanismos básicos para la identificación y reconocimiento de usuarios; otra funcionalidad es el seguimiento de las operaciones de los clientes (monitoreo).

## **Hardware y Software para Servidores**

La gran parte de los programas residentes en un servidor WWW, como ocurre con la mayoría del software para Internet, es de libre distribución. Los más populares son los servidores desarrollados en el CERN y NCSA, otros han sido desarrollados por grandes compañías como Netscape y Sun, los cuales impactan profundamente por sus avanzadas características de visualización, despliegue y facilidad de instalación.

Hasta hace algunos años, HTTPd de NCSA en sus distintas versiones, se localizaba a la vanguardia en distribución, seguido de servidores específicos para plataformas Macintosh (MachHTTP), y de HTTP del CERN.

En lo correspondiente al hardware, tradicionalmente Internet ha sido una extensión lógica de los sistemas con plataforma UNIX, se pone de manifiesto el predominio de este sistema operativo (por encima del 60% de los demás sistemas operativos) y en concreto, los equipos de Sun Microsystems Co., los cuales son líderes en el desarrollo de tecnología de hardware; aún así, los sistemas basados en computadoras personales (tanto PC como Macintosh) son una parte importante de la base instalada a nivel mundial.

Los servidores del CERN y NCSA fueron los más usados inicialmente, esto se atribuye a que fueron los pioneros en el área, además de ser de distribución gratuita, y ser sometidos a una continua evolución. En general se considera al servidor distribuido por el CERN como más seguro y completo, pero el de NCSA más rápido.

El software de servidores comerciales está generalmente más detallado y completo (suele incluir utilidades de visualización y control con análisis estadístico), también esta libre de fallas que presentan a veces los servidores de dominio público, con la gran ventaja de que al aparecer errores la respuesta es inmediata por parte del distribuidor, mientras que con el de dominio público, NO hay más solución que esperar a que la nueva versión resuelva el problema, lo mismo se aplica a la documentación y asesorías técnicas: *"cuando se paga por un software se dispone de garantía y de especialistas que ayudan en la solución de problemas que puedan surgir".*

## Software para Clientes

Este tipo de software son programas de aplicación que canalizan las peticiones del cliente hacia el servidor, e interpretan adecuadamente su respuesta. El número de estos programas crece continuamente, disponibles para un gran número de plataformas. Los más sencillos pueden ser usados en terminales sin resolución gráfica sin perder el carácter de hipertexto.

Gran parte de este software es de dominio público, sin embargo existe una compañía que acapara gran parte del mercado de visualizadores: **Netscape**, la cual tiene una política de difusión gratuita (proporcionando productos a prueba), y limitada en tiempo por nuevas versiones revisadas y liberadas (tanto de visualizadores-cliente como de servidores y editores HTML). Netscape ha conseguido una difusión que la sitúa como un estándar sin competencia, en una posición dominante en esta creciente industria.



Figura 2.4a. Visualizador Netscape.

Un gran número de usuarios (clientes), opta por el visualizador de Netscape, llamado Navigator. La trascendencia de este visualizador es tal que su particular versión del lenguaje HTML, así como sus extensiones propuestas, se usa ampliamente y afectan los trabajos de definición de nuevos estándares en otras compañías y organismos internacionales relacionados con el W3C.

Un pionero del hipertexto es Lynx, *visualizador en modo texto* que no hace uso del ratón, menús, ni alguna otra facilidad. Es simple, limitado y fácil de instalar, útil cuando no se dispone de un cliente gráfico. Este era el tipo de visualizadores habituales hasta la llegada de Mosaic:

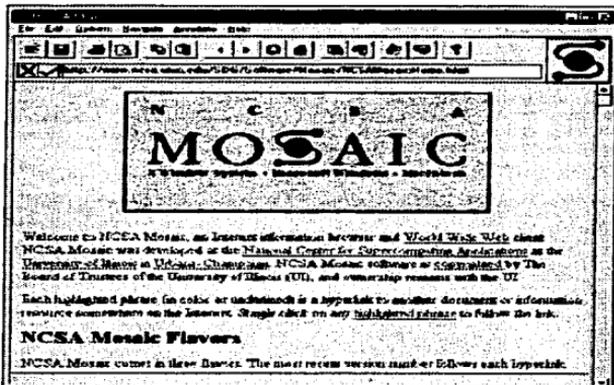


Figura 2.4b. Visualizador Mosaic.

Existen otros muchos visualizadores. Algunos tratan de unificar funciones de edición en HTML y visualización, otros explotan determinados aspectos de WWW, HTTP y TCP/IP.

La lista es muy extensa: Interworks, Macweb, WebExplorer (para OS/2), Cello, Arena, WinWeb, Chimera, MidasWWW, GWHIS, ViolaWWW o TkWWW.

Recientemente la compañía Microsoft se ha sumado a la lucha por el atractivo mercado millonario de visualizadores con su **Internet Explorer**. Se trata así como Netscape, de un software con muchas funcionalidades de presentación y modificaciones al lenguaje, asume buena parte de las extensiones propuestas por Netscape y añade algunas nuevas (se abordarán estas en el capítulo IV, Extensiones HTML).

## Capítulo III

# Lenguaje HTML

---

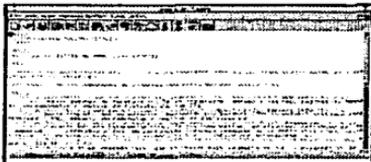
Las etiquetas HTML (lenguaje de marcación de hipertexto), se escriben en archivos planos, se puede emplear cualquier editor de texto, teniendo siempre en cuenta que se debe efectuar un respaldo del documento como un archivo de texto sin formato. Es posible emplear alguno de los numerosos convertidores de formato (disponibles en Internet), que pasan documentos de acuerdo con las especificaciones del programa a su equivalente HTML (xv para UNIX). Además hay editores diseñados para tal fin, es decir, programas en los que se escribe el contenido del documento HTML y se visualiza al mismo tiempo su resultado (WYSIWYG).

Existen programas de aplicación como: HTML Assistant, el cual es un editor de libre distribución que depende de un visualizador externo para ver el trabajo, HTMLed, herramienta para Windows, WebAuthor, aplicación comercial que se apoya en Word de Microsoft y convierte los documentos HTML a formato con la extensión .doc, HoTDog, herramienta comercial que trabaja sobre Windows y en ciertas versiones de UNIX (muy solicitada), existen muchas más como: Wedit, tkHTML, tkWWW, HtmlText, HTMLed, etc.

En general cualquiera de estas aplicaciones sirven, sin embargo, las que pueden ser de mayor utilidad son aquellas que trabajan con un editor diseñado para tal fin, que permita insertar etiquetas HTML y deje plena libertad para incluir cualquier texto, además de permitir una visualización plena.



XV. Visualizador/convertor en UNIX



HotDog. Editor HTML en Windows.

Figura 3a. Ejemplos de herramientas para la edición de páginas.

### 3.1. Etiquetas del lenguaje HTML

**HTML** (*HyperText Markup Language*), es un lenguaje de marcación de hipertexto, se trata de un lenguaje empleado para marcar o señalar las partes estructurales de un documento, (permitiendo así, la definición de las distintas instrucciones y su contenido). Estas marcas se denominan **etiquetas** y son la base principal del lenguaje HTML. Así un documento HTML es un archivo de texto con etiquetas las cuales proporcionan la apariencia final del documento a través de un cliente o visualizador específico.

#### **Niveles HTML:**

HTML deriva del estándar **SGML** (**Standard Generalized Markup Language**) o Lenguaje Normal Generalizado de Marcación (norma ISO). El objetivo principal que persigue SGML es colaborar en el formato de información en línea de manera que su distribución, búsqueda y recuperación electrónicas sean independientes de los detalles de la apariencia final del documento. SGML se emplea para definir HTML en todos sus niveles.

Actualmente existen cuatro niveles HTML: 0, 1, 2 y 3. Los *niveles 0 y 1* son un subconjunto de las instrucciones totales de HTML que **todos** los visualizadores entienden (sean gráficos o basados en texto), sin embargo están limitados en poder de comunicación y despliegue. El *nivel 2* comprende todas las instrucciones de los niveles 0 y 1, extensiones de algunas de ellas, y la inclusión de otras nuevas. El *nivel 3* es el más reciente, y comprende al igual que el nivel 2, todas las instrucciones de sus antecesores, ampliaciones a algunas de ellas, y la incorporación de nuevas capacidades.

Los niveles HTML pueden verse como versiones del lenguaje: *cada nuevo nivel contempla a su antecesor y añade características adicionales, basadas éstas en avances tecnológicos tanto de hardware y software.*

Al considerar los diferentes niveles de HTML, un desarrollador tiene la certidumbre de saber cuáles elementos de su aplicación podrán ser mostrados a través del Web e Internet, en cualquier visualizador, sin problema y cuáles no (sirviendo esto para una buena implementación).

#### **Etiquetas:**

Una **etiqueta** es un texto incluido entre los símbolos menor que <, y mayor que >. El texto incluido dentro de estos símbolos es explicativo de la utilidad de la etiqueta. Por ejemplo:

<B>	Letra resaltada, (del inglés Bold).
<TABLE>	Define una tabla.
<IMG>	Incluye una IMaGen.

Existe normalmente una etiqueta de inicio y otra de fin; la de fin contiene el mismo texto que la de inicio añadiéndole al principio una barra inclinada /. El efecto que define la etiqueta tiene validez para todo lo que este incluido entre las etiquetas de inicio y fin, ya sea texto plano u otras etiquetas HTML.

<ETIQUETA> elementos afectados por la etiqueta </ETIQUETA>

Por ejemplo, con la etiqueta siguiente:

```
<B>Texto resaltado</B>.
```

Obtendremos:

**Texto resaltado**

Algunas etiquetas no necesitan terminar con /, son aquellas en las que el final este implícito, por ejemplo <P> párrafo, <BR> salto de línea ó <IMG> inclusión de una imagen. Estas definen un efecto que se produce en un punto determinado sin afectar a otros elementos. El uso de mayúsculas o minúsculas en las etiquetas es indiferente, se interpretan del mismo modo en ambos casos, pero lo común es expresarias en mayúsculas para que destaquen con más nitidez del texto normal.

### **Atributos de las etiquetas:**

Las etiquetas pueden presentar modificadores llamadas **atributos** que permiten definir diferentes posibilidades de la instrucción HTML. Estos atributos se definen en la etiqueta de inicio y consisten normalmente en el nombre del atributo y el valor que toma, separados por un signo de igual. El orden en que se incluyen los atributos es indiferente, no afectando al resultado. Si se incluyen varias veces el mismo atributo con distintos valores el resultado obtenido será imprevisible dependiendo de la interpretación del visualizador. Cuando el valor que toma el atributo tiene más de una palabra deberá expresarse entre comillas, en otro caso no será necesario. Un ejemplo de atributo es:

```
<A HREF="http://www.unam.mx">Pagina principal de la UNAM</A>
```

En este caso la etiqueta A presenta un atributo HREF cuyo valor es el URL: <http://www.unam.mx>

Igualmente una etiqueta puede presentar varios atributos:

```
<HR ALIGN=LEFT NOSHADE SIZE=5 WIDTH=50%>
```

En este caso la etiqueta HR presenta cuatro atributos. El segundo atributo *NOSHADE* es un caso especial que no presenta valor. El orden en que se especifiquen los atributos NO afecta el resultado final.

### **Etiquetas correctas:**

Todo texto que se encuentre entre los caracteres < y > se considera una etiqueta, si esta no es una de las válidas en HTML no es tomada en cuenta, sin causar algún tipo de error. Dejándose el texto o las etiquetas a las que afectaba como si no existiera la etiqueta extraña. Cuando se comete un error sintáctico, al expresar una etiqueta o un atributo no se produce error, simplemente no se obtiene el efecto deseado.

El lenguaje HTML es un lenguaje que evoluciona rápidamente y cada nueva versión de programas visualizadores ó clientes presentan etiquetas nuevas que causan efectos espectaculares o atributos nuevos de las etiquetas ya existentes. Esto causa que las aplicaciones más antiguas no entiendan estas nuevas etiquetas y por tanto las considere erróneas y no realicen la acción esperada. Dándose el caso de atributos que son válidos para un solo visualizador.

Cuando se crea código HTML hay que hacerlo lo más estándar posible para permitir que el documento pueda ser visto de forma efectiva por distintos clientes en diversas computadoras. Por tanto se debe renunciar a efectos espectaculares que sólo tengan validez en un visualizador, e intentar comprobar como se ve el documento en una variedad de clientes.

## **3.2. Formato URL**

URL (Uniform Resource Locator), significa localizador uniforme de recursos y permite acceder a cualquier recurso en la red desde el visualizador Web. WWW pretende unificar el acceso a información de variados servicios y aplicaciones, que son incompatibles entre sí; de esta manera, en una misma aplicación-cliente se puede acceder a recursos en forma uniforme y a servicios tales como FTP, gopher, WAIS, etc., permitiendo que los documentos HTML incluyan enlaces a diversas fuentes de información.

### **Uso y Formato:**

Los URL se utilizan para definir el documento destino de los hiperenlaces, para referenciar los gráficos y cualquier otro archivo que se desee incluir dentro de un documento HTML. Cada elemento de Internet tiene un URL que lo define, ya que se encuentra en un servidor de WWW, FTP, News, etc.

El formato es:

**servicio://máquina.dominio:puerto/ruta/archivo**

Donde, **servicio** puede ser:

- **http:** (HyperText Transport Protocol), es el protocolo utilizado para transmitir hipertexto. Todas las páginas HTML en servidores WWW deben ser referenciadas mediante este servicio, este indica conexión a un servidor Web.
- **https:** (HyperText Transport Protocol Secure), es el protocolo para la conexión a servidores WWW "seguros". Estos servidores son normalmente de ámbito comercial y utilizan encriptación de datos para evitar la interceptación de estos durante una transmisión, usualmente números de tarjeta de crédito, datos confidenciales, etc.
- **ftp:** (File Transfer Protocol), emplea el protocolo FTP de transferencia de archivos. Se utiliza cuando la información que se quiere acceder se encuentre en un servidor de ftp. Por omisión se accede a un servidor anónimo (anonymous), si se desea indicar el nombre de usuario se usa: `ftp://máquina.dominio@usuario`, y luego se pedirá la clave de acceso.
- **gopher, wais:** Cualquiera de estos servicios de localización de información, hace referencia a diversos directorios para localizar información.
- **news:** Proporciona el servicio de noticias *news*, para ello el visualizador Web debe ser capaz de proporcionarlo. Se hace referencia al servidor *news* y como ruta el grupo de noticias al que se desea acceder, por ejemplo: `news://news.java.com/`
- **telnet:** Emulación de terminal remota, comunemente usado para acceder a cuentas de dominio público. Lo normal es efectuar la llamada a una aplicación externa que realice la conexión, en este caso se indica la máquina y el login: `telnet://máquina.dominio@login`
- **mailto:** Se emplea para enviar correo electrónico, no todos los visualizadores tienen esta capacidad. En este caso solo se indica la dirección de correo electrónico del destino: `mailto://ulises@morsa.dgsca.unam.mx`

La **máquina.dominio** indica el servidor que nos proporciona el recurso, en este caso se emplea el protocolo IP para identificar el nombre de la computadora y el dominio. En el caso de nuestra Universidad el dominio siempre será `unam.mx`. Por lo tanto un nombre válido es: `serpiente.dgsca.unam.mx` ó `www.unam.mx`.

Es muy importante indicar siempre el dominio, ya que se debe suponer que se conectarán a nuestras páginas desde servidores externos a la red local, por tanto, si no se indica el dominio, los URL especificados no podrán ser seguidos por visualizadores externos.

El **puerto TCP** es opcional y lo normal es no colocarlo, si el puerto es el mismo que es empleado normalmente por el servicio. Solo se especifica cuando el servidor emplea un puerto distinto al puerto por omisión.

El camino es la **ruta** de directorios que hay que seguir para encontrar el documento que se desea referenciar. En el caso de UNIX para separar los subdirectorios se emplea la

barra /, por conveniencia, al ser este tipo de computadoras las más usadas como servidores Web. El nombre de los subdirectorios y del archivo referenciado puede ser de más de ocho caracteres y se toma muy en cuenta la diferencia entre mayúsculas y minúsculas.

La extensión de los archivos es importante, así el servidor reconoce el tipo de documento que se accede e indica al cliente el modo en que debe tratarse ese documento. Para definir los tipos de documentos se emplean los tipos MIME. Las extensiones más normales con sus tipos correspondientes son:

Tipo MIME	Extensión	Tipo de archivo
text/html	.html ó .htm.	documento HTML
text/plain	.txt	por omisión, texto plano
image/gif	.gif	imagen de formato GIF
image/jpeg	.jpg ó .jpeg	imagen de formato JPEG

El visualizador WWW, realiza una acción para cada tipo de archivo, solo los del tipo **text/html** serán mostrados como documentos HTML. En el caso de que el tipo no sea reconocido por el cliente, se considerará (por omisión) como un documento de texto normal.

Si no se indica un archivo y solo se referencia un directorio, se accede a la página html por omisión de ese directorio. En el servidor Web se encuentran definidos ciertos archivos para ser usados (*default.htm* ó *default.html*). En caso de que no exista este archivo, se muestra un listado de todos los documentos que forman el directorio. Este archivo debe ser la página inicial (home page) del servidor o del espacio Web asignado en el disco duro de la computadora. Algunos **ejemplos de URL** son:

URL	Descripción
<a href="http://informatica.aragon.unam.mx">http://informatica.aragon.unam.mx</a>	En este caso solo se indica el tipo de servicio y la maquina con dominio. El resto de los parámetros se toman por omisión (el puerto 80, el directorio, la raíz del servidor y el documento por omisión de ese directorio).
<a href="http://www.unam.mx/temas/temas.html">http://www.unam.mx/temas/temas.html</a>	Este URL es más completo en este caso se accede al archivo temas.html que se encuentra en el directorio temas del servidor de WWW, <a href="http://www.unam.mx">www.unam.mx</a>
<a href="ftp://ftp.playboy.com/playmates/sep.playmates.html">ftp://ftp.playboy.com/playmates/sep.playmates.html</a>	En este caso se accederá a un servidor de FTP anónimo: <a href="ftp://ftp.playboy.com">ftp.playboy.com</a> , por el protocolo FTP y se accederá al archivo <a href="ftp://ftp.playboy.com/playmates/sep.playmates.html">sep.playmates.html</a> del directorio <a href="ftp://ftp.playboy.com/playmates/">mates</a> .
<a href="mailto://ccc@polaris.labvis.unam.mx">mailto://ccc@polaris.labvis.unam.mx</a>	Enviará un mail al equipo del Laboratorio de Visualización de la DGSCA en la UNAM (opción soportada por visualizadores modernos).

### 3.3. Estructura de un documento HTML

Un documento HTML está definido por una etiqueta de apertura `<HTML>` y una etiqueta de cierre `</HTML>`. Entre estas hay dos partes fundamentales: el *encabezado*, delimitada por la etiqueta `<HEAD>` y el *cuerpo*, delimitado por la etiqueta `<BODY>`. Por tanto la estructura básica de un documento HTML es:

```

<HTML>
  <HEAD>
    Definiciones de el encabezado
  </HEAD>
  <BODY>
    Instrucciones HTML
  </BODY>
</HTML>

```

Ninguno de estos elementos es obligatorio, se puede crear código HTML que no muestren algún problema sin incluir estas etiquetas de identificación. Si se utilizan elementos que forzosamente deban ser incluidos en el encabezado (como la etiqueta de título), no son reconocidos correctamente si no se incluyen entre las etiquetas de `<HEAD>`.

#### Comentarios:

Para insertar comentarios dentro de un documento HTML se emplea la etiqueta: `<!--`, definiéndose un comentario de la forma:

```
<!-- Esto es un comentario -->
```

Los comentarios son útiles para la identificación del documento, (título, autor, fecha en el que fue realizado, etc.), empleándola antes de la etiqueta `<HTML>`. Los comentarios no se muestran a través del visualizador, el único modo de observarlos es viendo el código fuente del archivo (opción disponible en algunos clientes Web).

### 3.4. Encabezado de un documento HTML

El encabezado de un documento HTML está delimitado por las etiquetas `<HEAD>` y `</HEAD>`, aquí se incluyen las definiciones generales que afectan a todo el documento.

La etiqueta `TITLE` tiene un uso general y aunque es opcional se recomienda incluirla en todos los documentos a crear. Los distintos componentes que pueden formar el encabezado de un documento HTML son: `<TITLE>`, `<BASE>`, `<INDEX>`, `<META>`.

**<TITLE> título del documento:**

Especifica el título del documento, todo documento HTML debe tener un título. Este debe guardar relación con el contenido del documento y definirlo de forma general, su tamaño no es limitado (es recomendable que no sea extenso). Dentro de esta etiqueta no se puede usar ninguna de las restantes.

El título no forma parte del documento en sí, normalmente se observa en la parte superior de la ventana de la aplicación cliente. Se utiliza principalmente para identificar la página en los marcadores (*bookmarks*) y la lista de documentos accedidos (*history list*), también se emplea por algunos servidores como resultado de una búsqueda para poder intuir el contenido del documento.

Esta etiqueta, debe ser usada dentro de las etiquetas que definen el encabezado de la siguiente forma:

```
<HEAD>
  <TITLE>Título del documento</TITLE>
</HEAD>
```

**<BASE> URL base del documento:**

Especifica el URL que se toma como base del documento HTML, esta se emplea para las referencias a *URL relativas*, que se encuentran en los hiperenlaces y en las referencias de las imágenes. Comúnmente es utilizada en documentos obtenidos de otros servidores o directorios e incluidos fuera de contexto, de esta forma los enlaces siguen siendo válidos. Tiene un atributo **HREF** que indica el URL base del documento, el formato es:

```
<BASE HREF="URL">
```

Por ejemplo:

```
<BASE HREF="http://serpiente.dgsca.unam.mx">
```

**<ISINDEX> documento índice:**

Indica que el documento es un índice, y por tanto se puede realizar una búsqueda. Presenta un indicador preguntando la *palabra clave* de la búsqueda. Se emplea en documentos realizados con CGI (Common Gateway Interface). Presenta dos atributos, **PROMPT** que indica el texto que aparece como inductor de la búsqueda y **ACTION** que especifica el archivo cgi que trata la consulta, por omisión será el documento actual.

**<META> llamada automática del documento:**

Su principal función es indicar documentos con *refresh* automático. (Indica un documento que debe sustituir al actual en un determinado número de segundos). El formato es:

```
<META HTTP-EQUIV="REFRESH" CONTENT="número_segundos
;URL=URL_de_refresco">
```

Se indica el número de segundos que deben pasar antes del refresco y el documento HTML que sustituye al actual. Si se indica cero segundos, la transición entre uno y otro documento será inmediata. Si no se indica el URL el documento actual se refrescará. Una utilidad es para documentos que cambien con una gran frecuencia y transcurrido un tiempo se actualicen, o documentos que deban ser vistos de forma secuencial, se empieza por un documento pasando a los siguientes transcurrido un periodo de tiempo sin necesidad de acción alguna por parte del lector.

### 3.5. Cuerpo de un documento HTML

El cuerpo de un documento HTML está delimitado por `<BODY>` y `</BODY>` y en él se incluyen todas las líneas de código HTML y el texto que forman el documento, es similar al *BEGIN* ó *¿*, de un lenguaje de programación. Al igual que el encabezado (*HEAD*) es opcional, pero es recomendable para la identificación de las distintas zonas del documento. Si un documento no presenta ninguna de las etiquetas de identificación de sus distintas partes (`<HTML>`, `<HEAD>` ó `<BODY>`) se considera que todo lo que se defina pertenece al cuerpo del documento.

#### Definición de colores:

Para dar color al texto, se emplea el formato: `#rrvvaa`.

Donde se indica en hexadecimal la proporción de rojo, verde y azul que forma el color deseado (por ser estos los colores básicos del espectro de luz natural). El símbolo `#` es opcional. Así el número `0F` será el 15, `0E` será 14, etc. En la especificación del color se utiliza para definir la proporción de cada color un número del 0 al `FF` (255), 0 indica nada de ese color y `FF` la mayor proporción del color.

Algunos ejemplos son:

#000000	Negro
#FFFFFF	Bianco
#FF0000	Rojo
#00FF00	Verde
#0000FF	Azul

Algunos colores se encuentran predefinidos y pueden ser referenciados por su nombre y son válidos únicamente para 2 visualizadores: *Netscape Navigator* y *Microsoft Internet Explorer*, estos son: *black, olive, teal, re, blue, maroon, navy, gray, lime, fuchsia, white, green, purple, silver, yellow, aqua*.

### Atributos de <BODY>

La etiqueta **BODY** presenta algunos atributos que son de definición global para todo el documento, estos definen los colores y el fondo del documento HTML. Los atributos de **BODY** son:

**<BODY BACKGROUND="#URL" BGCOLOR=#mvvaa TEXT=#mvvaa LINK=#mvvaa VLINK=#mvvaa >**

- **BACKGROUND="#URL":**  
Define la imagen gráfica que se utiliza de fondo en el documento HTML, el URL definido es la ruta a un gráfico. Este se muestra como fondo del texto. En el caso de que la imagen no llene todo el fondo del documento esta es reproducida tantas veces como sea necesario.
- **BGCOLOR=#mvvaa** ó nombre\_del\_color:  
Indica el color del fondo del documento, solo se emplea si no se ha definido una imagen de fondo, o si esta no puede obtenerse.
- **TEXT=#mvvaa** ó nombre del color:  
Especifica el color del texto normal dentro del documento HTML. Por omisión es negro.
- **LINK=#mvvaa** ó nombre del color:  
Indica el color que tendrán los hiperenlaces que no han sido accedados. Por omisión es azul.
- **VLINK=#mvvaa** ó nombre del color:  
Color de los enlaces que ya han sido visitados. Por omisión es de color azul oscuro.

Algunos ejemplos son:

Código HTML	Descripción
<code>&lt;body background="http://www.unam.mx/fondos/arena.gif"&gt;</code>	El documento tiene como fondo la imagen indicada en el URL.
<code>&lt;body bgcolor=blue text=white link=red vlink=aqua&gt;</code>	El fondo será azul, el texto blanco, los enlaces rojos y los enlaces visitados color <i>aqua</i> .

### 3.5.1. Espaciados y saltos de línea

El código HTML no permite más de un elemento blanco (espacios, tabuladores, saltos de línea, etc.), separando cualquier elemento o texto, todos estos son convertidos a un único espacio blanco y el resto se omiten en la representación del documento. En el documento fuente podremos usar el espaciado que deseemos. Existen etiquetas especiales para definir estos elementos de control de texto:

#### **<P> cambio de párrafo:**

Define un párrafo, se usa al principio o final de un párrafo de texto e introduce un espaciado de separación (normalmente dos líneas) con el próximo texto que se exprese. Esta etiqueta se puede utilizar para introducir un espaciado entre cualquier elemento HTML. El efecto se consigue introduciendo la etiqueta <P> en el punto en el que se desee. La etiqueta de fin de párrafo </P> es opcional no siendo necesario incluirla. Existen elementos HTML que incluyen separaciones de líneas, tanto antes como después, en estos casos no es necesario, pero posible, introducir el elemento de párrafo.

No es necesario utilizar esta etiqueta ni antes ni después de encabezados <Hn>, después de <HR> (reglas horizontales), <ADDRESS>, <BLOCKQUOTE>, <PRE>. Tampoco es necesario dentro de <LI>, <DT> ni <DD>, que se utilizan para separar los distintos elementos de una lista.

#### **<BR> salto de línea:**

Similar al anterior pero en este caso el espaciado del texto es menor, se pasará a la línea siguiente, sin dejar una línea de separación. En este caso es un cambio de línea y no de párrafo. Igualmente no es necesario usarlo después de los elementos que llevan implícitos un salto, ni tampoco es necesaria la etiqueta de fin </BR>.

#### **<HR> margen horizontal:**

**<HR ALIGN=LEFT|RIGHT|CENTER NOSHADE SIZE=n WIDTH=n>**

Se emplea para dividir un documento en distintas secciones, muestra una línea horizontal de tamaño determinable. Se asemeja al salto de página dentro de un documento. Si no se especifican atributos, dibuja una línea que ocupa el ancho de la pantalla e introduce una separación con el texto anterior y siguiente, equivalente al cambio de párrafo. No es necesaria la etiqueta de fin </HR>. Con los atributos podemos especificar su forma y tamaño.

Esto atributos son:

- **ALIGN=LEFT, RIGTH ó CENTER:**  
Indica la forma en la que se alinea la regla en caso de no ocupar todo el ancho de la pantalla del visualizador, a la izquierda, derecha o centrada.
- **NOSHADE:**  
No muestra la sombra de la barra, evitando el efecto en tres dimensiones.
- **SIZE=n:**  
Indica la altura de la regla en puntos de la pantalla.
- **WIDTH=n ó n%:**  
Especifica el ancho de la regla, se puede especificar en tanto por ciento del ancho de la ventana `<HR WIDTH=50%>` o en valor absoluto `<HR WIDTH=75>` en puntos de la pantalla.

#### **<PRE> texto preformateado:**

Muestra un porción de texto en el que se respetan los saltos de línea, tabuladores y espacios en blanco. Todo lo que se encuentre entre las etiquetas de inicio y fin de texto preformateado se muestra tal y como se expresa en el código fuente del documento HTML. Para mostrar este texto se utiliza una fuente de espaciado fijo más pequeña que el texto normal. Este presenta un problema cuando la línea es demasiado larga, ya que no la corta el visualizador y se muestra, tal y como aparece en el archivo fuente. No se recomienda como forma rápida de convertir un documento texto a formato visible, ya que presenta limitaciones y no se obtiene una buena presentación. Dentro de la etiqueta PRE no tienen validez la mayoría de las instrucciones HTML, solo se podrán utilizar **<B>**, **<I>** y los hiperenlaces. El resto de las etiquetas son ignoradas.

#### **<CENTER> centrado de texto e imágenes:**

Se utiliza para centrar líneas de texto, imágenes o cualquier otro elemento de código (tablas, listas, etc.). Todo lo que se encuentre entre las etiquetas de inicio y fin aparecerá centrado en el programa cliente. Por ejemplo:

`<CENTER>` Este texto ó imagen se verá centrado `</CENTER>`

El texto aparece centrado en el visualizador. Se puede utilizar tanto para una línea como para un párrafo o imagen, y para cualquiera de las instrucciones del lenguaje.

#### **&nbsp; (espacios en blanco):**

Con esta secuencia de caracteres se logran espacios en blanco, pudiendo mostrar más de un espacio en blanco de separación. Se deben incluir tantas expresiones **&nbsp;** como espacios en blanco se deseen conseguir.

### 3.5.2. Caracteres latinos y especiales

Los caracteres acentuados y algunos caracteres especiales que usa el lenguaje HTML para definir sus etiquetas no se pueden incluir en un documento de manera normal, se deben utilizar una serie de *secuencias de escape* que al desplegar el documento se sustituyen por el carácter deseado.

Estas *secuencias de escape* comienzan todas con el símbolo *ampersand* (&), seguido de un texto (siempre en minúsculas) que define el carácter deseado y termina con el símbolo punto y coma (;). El error más común cuando se usan estas secuencias de escape es no utilizar el punto y coma final, en cuyo caso se muestra el literal que define la secuencia, en lugar del carácter deseado. No es necesario dejar espacios en blanco ni antes ni después de los caracteres especiales para que queden perfectamente acomodados en la palabra.

Al incluir en un documento HTML el símbolo *menor que* (<) o *mayor que* (>) se interpretará siempre como la definición de una etiqueta y por tanto no se muestran al interpretar el documento. Para expresar estos símbolos y otros propios del lenguaje HTML, se emplean ciertas secuencias de escape:

Sec. Escape	Símbolo
&lt;	Signo < (menor que)
&gt;	Signo > (mayor que)
&amp;	Signo & (ampersand)
&quot;	Se muestra el signo de comillas ("")

#### Caracteres acentuados:

Existen una serie de etiquetas que permiten mostrar caracteres acentuados y caracteres latinos (como: ñ). Estos caracteres al incluirse en un documento HTML sin utilizar las secuencias de escape, se muestran correctamente, el visualizador podrá leer estos caracteres correctamente siempre y cuando tenga el mismo código de caracteres (español) que el autor del documento instalado en la computadora.

Los primeros 127 caracteres del código ASCII son comunes para todos los países e incluyen todas las letras del alfabeto, a partir del 128 son específicos para cada lenguaje, por tanto si alguien desde otro país que tenga un código incompatible intenta leer el documento, probablemente encuentre caracteres extraños y por tanto el cliente no será capaz de leer los caracteres acentuados.

Existen diversas secuencias que definen los distintos tipos de acentos: agudo, grave o circunflejo. Para el acento agudo se emplea el literal *acute*, tanto para las mayúsculas como para las minúsculas. De esta manera se incluye el símbolo de ampersand (&),

seguida de la vocal que deseamos acentuar, la palabra acute y finalmente el símbolo punto y coma (;). Representándose los acentos de la siguiente manera:

Sec. Escape	Letra	Sec. Escape	Letra
&aaacute;	á	&Aacute;	Á
&eaacute;	é	&Eacute;	É
&iaacute;	í	&Iacute;	Í
&oaacute;	ó	&Oacute;	Ó
&uaacute;	ú	&Uacute;	Ú

Para la letra ñ se emplea la secuencia tild: &ntilde; ó &Ntilde;

### Simbolos:

Para expresar una carácter por su valor en ASCII, se emplea el símbolo #, seguido de su equivalente numérico. Para el acento circunflejo el literal *circ* y para la diéresis: *uml*. Para expresar los símbolos de interrogación y admiración:

Sec. Escape	Letra	Sec. Escape	Letra
&#191;	¿	&#161;	¡
&uuml;	ü	&Uuml;	Ü
&icirc;	ï	&Icirc;	Ï

### 3.5.3. Encabezados

#### <H1> - <H6> encabezado de títulos:

En HTML es posible definir seis tipos distintos de encabezados que son normalmente el título del documento y los diversos apartados que lo integran. Las etiquetas que definen los encabezados son: <H1>, <H2>, <H3>, <H4>, <H5>, <H6>. El texto indicado entre las etiquetas de inicio y de fin es el afectado por los encabezados:

Para la instrucción:

```
<H1>Este texto aparecerá resaltado</H1>
```

el resultado será:

**Este texto aparecerá resaltado**

Lo normal es emplear los encabezados en forma consecutiva y descendente, es decir, comenzando el documento con el encabezado `<H1>` que define el título, luego para los subtítulos principales se emplea `<H2>` ó `<H3>`, etc.

Sin embargo se puede emplear cualquier encabezado en el punto que se desee para resaltar cualquier texto. El modo de presentación pueden variar de un programa cliente a otro.

#### **`<Hn ALIGN=CENTER>` Alineación del encabezado**

Puede presentar un atributo, que especifica que el encabezado se muestre centrado. Por ejemplo: `<H1 ALIGN=CENTER>`Este texto aparecerá resaltado y centrado`</H1>`

### **3.5.4. Tamaño y Color de las fuentes de caracteres**

Existe a parte de los encabezados que permite una más sencilla adaptación del tamaño de las fuentes, una etiqueta para modificar su color. Así es posible incluir texto resaltado en medio de una frase (con los encabezados no lo es), ya que estos introducen automáticamente un salto de línea consecutivo. La etiqueta es `FONT` y presenta atributos que permiten modificar el tamaño y color del texto incluido entre la etiqueta de inicio y fin.

#### **`<FONT SIZE=n>` tamaño de la fuente:**

El atributo `SIZE` permite indicar el tamaño de la fuente, su valor puede estar entre 1 y 7. Incrementándose de tamaño progresivamente desde 1, hasta 7, fuente de mayor tamaño. El texto normal es equivalente a la fuente 3. El tamaño también puede indicarse en forma relativa, indicando el incremento o decremento a partir del tamaño base. Por tanto si se indica como valor +1 la fuente será de 4. Existe una etiqueta que redefine la fuente por omisión, esta etiqueta es:

`<BASEFONT SIZE ...>`

Los elementos de tamaño de fuente pueden ser definidos para todo un documento, teniendo validez dentro de elementos tales como listas y formas, pero no tendrán validez global en las tablas, debiendo definir cada una de las celdas al tamaño de fuente deseado.

Algunos ejemplos de tamaños de fuentes son:

Código HTML	Descripción
<FONT SIZE=2>UNAM</FONT>	Se obtiene texto de tamaño de fuente 2.
<FONT SIZE=-1>UNAM</FONT>	Se obtiene texto de tamaño de fuente 2 (en forma relativa: 3-1=2).
<FONT SIZE=+2>UNAM</FONT>	Fuente de tamaño 5 (relativa: 3+2=5).
<FONT SIZE=6>UNAM<FONT SIZE=1> ENEP</FONT>ARAGON</FONT>	Se pueden combinar varios tamaños.
<FONT SIZE=4> ENEP Aragon <B> </> <A HREF="http://informatica.arogon.unam.mx"></A> UNAM</> Nuestro futuro HOY </B> </FONT>	Se puede combinar con elementos de resalte e hiperenlaces.

**<BASEFONT SIZE=n> fuente por omisión:**

Esta proporciona el tamaño de la fuente que se considera como base para definir los tamaños de fuente relativos. Esta etiqueta no define el tamaño de la fuente por omisión, para el texto normal, y solo se emplea para cálculos de tamaño relativos. Se puede usar cuando en un documento en el que todos los tamaños estén definidos de forma relativa, se desee cambiar de forma global el tamaño. Al no incluir esta etiqueta el valor base para estos cálculos, por omisión es 3.

**<FONT COLOR=texto color ó rrvvas> color de la fuente:**

El atributo COLOR nos permite definir el color del texto incluido entre las etiquetas de inicio y fin. Este solo funciona para *Internet Explorer* de Microsoft y *Netscape Navigator* 2.0 o superior. Este atributo puede ser definido en forma conjunta con SIZE, o independientemente. Si se define el atributo SIZE solo, el color del texto será tomado por omisión, si se define solo con la etiqueta COLOR, el tamaño será el de la fuente base. Al igual que en el caso del tamaño de la fuente, pueden ser utilizados junto con otros elementos del lenguaje HTML como listas y formas, además de elementos de resalte. El cambio de color no afecta al texto o elementos incluidos en un hiperenlace, empleándose, según sea el caso, el color por omisión definido en la etiqueta BODY. Algunos ejemplos:

Código HTML	Descripción
<FONT COLOR=FF00FF>DGSCA</FONT>	El texto se muestra de un color lila, al mezclar los colores rojo y azul.
<FONT COLOR=AQUA>DGSCA</FONT>	También se puede utilizar uno de los colores predefinidos.
<FONT SIZE=5 COLOR=red> DGSCA</FONT>	Se puede combinar con el tamaño de la fuente.
<FONT COLOR=NAVY>DG<FONT SIZE=4 COLOR=BLUE>SC</FONT>A</FONT>	Se combinan varios colores en una misma frase o palabra.
<H3> <FONT COLOR=81426E> DG</>SC</>A</FONT> </H3>	Pueden ser utilizados junto con otros elementos de resalte como encabezados ó estilos físicos y lógicos.

### 3.5.5. Estilos físicos y estilos lógicos

Como estilos se definen los distintos efectos que se pueden aplicar al texto normal. Estos efectos son los resaltes (resaltado, cursiva, subrayado, etc. ...) que se pueden usar.

En HTML existen dos grupos principales de estilos: los lógicos y los físicos. Los físicos son aquellos que siempre causan un mismo efecto (resaltado, cursiva, ...) y los lógicos indicarán un tipo de texto (cita, nombre de persona, ...) que por sus características tiene un modo de mostrarse propio. Se pueden emplear ambos estilos para especificar un mismo efecto, por ejemplo para obtener una frase resaltada se puede usar el estilo físico (resaltado) o el estilo lógico (STRONG) y en ambos casos el resultado es el mismo.

Es recomendable utilizar el estilo lógico, ya que el modo en que se ven los distintos efectos puede ser definido por el usuario, a manera de obtener una presentación personalizada.

En un futuro los visualizadores podrán añadir modos más sofisticados de presentar los estilos y si se encuentran definidos de forma lógica la adaptación será más sencilla. Al utilizar estilos lógicos el modo de escribir código HTML se hace independiente de como se presenta finalmente el texto.

La tendencia actual de los visualizadores es el uso de los estilos físicos, olvidando un poco la versatilidad de los lógicos, sin embargo ambos estilos pueden ser usados indistintamente, sin ningún problema.

#### ***Estilos Físicos:***

El efecto se aplica al texto expresado entre la etiquetas de inicio y fin, estos son:

Etiqueta HTML	Descripción
<B>	Resaltado.
<I>	Cursiva.
<TT>	Máquina de escribir, muestra una fuente de caracteres de espaciado fijo.
<BLINK>	Parpadeo.
<SUB>	Subíndice. Para Netscape 2.0+
<SUP>	Super índice. Para Netscape 2.0+
<BIG>	Texto grande, se utilizará el mayor tamaño de fuente. Para Netscape 2.0+
<SMALL>	Texto pequeño, se utilizará la fuente de menor tamaño. Para Netscape 2.0+

### Estilos lógicos:

Etiqueta HTML	Descripción
<ADDRESS>	Se utiliza para especificar direcciones de correo electrónico.
<BLOCKQUOTE>	Introduce citas textuales o texto destacado, en este caso el texto se separa del texto circundante y se ajusta a los márgenes derecho e izquierdo.
<CITE>	Indica el título de una película o un libro.
<CODE>	Código fuente de un lenguaje de programación.
<DFN>	Especifica una definición.
<EM>	Indica énfasis.
<KBD>	Texto introducido desde el teclado
<SAMP>	Mensajes de estado de la computadora.
<STRIKE>	Texto desechado. (tachado)
<STRONG>	Especifica texto resaltado.
<VAR>	Indica una variable.

### 3.5.6. Listas

Existen tres tipos de listas: numeradas, sin numerar y definición. Las listas numeradas representan los elementos de la lista, numerando cada uno de ellos según el lugar que ocupan. Para este tipo de lista se emplea la etiqueta <OL></OL>. Cada uno de los elementos de la lista irá precedido de la etiqueta <LI> <OL>, esta puede llevar los siguientes atributos:

**start** = num

Indica que número es el primero de la lista. Si no se indica, empezará por el número 1.

**type** = tipo

Indica el tipo de numeración utilizada. Si no se indica será una lista ordenada numéricamente.

Los tipos posibles son:

- 1 = Numéricos. (1,2,3,4,...)
- a = Letras minúsculas. (a,b,c,d,...)
- A = Letras mayúsculas. (A,B,C,D,...)
- i = Números romanos en minúsculas. (i,ii,iii,iv,v,...)
- I = Números romanos en mayúsculas. (I,II,III,IV,V,...)

Código HTML	Resultado
<pre>&lt;OL&gt; &lt;LI&gt;Mexico &lt;LI&gt;Canada &lt;LI&gt;Brazil &lt;LI&gt;Chile &lt;/OL&gt;</pre>	1.Mexico 2.Canada 3.Brazil 4.Chile
<pre>&lt;OL type = A &gt; &lt;LI&gt;Mexico &lt;LI&gt;Canada &lt;LI&gt;Brazil &lt;LI&gt;Chile &lt;/OL&gt;</pre>	A.Mexico B.Canada C.Brazil D.Chile

Las listas sin numerar representan los elementos de la lista con una marca que antecede a cada uno de ellos. Se utiliza la etiqueta `<UL></UL>` para delimitar la lista, y `<LI>` para indicar cada uno de los elementos. La etiqueta `<UL>` puede contener el atributo **type** que indica la forma de la marca que antecede a cada elemento de la lista. Los valores de **type** pueden ser **disk**, **circle** o **square**, con lo que la marca puede ser un disco, un círculo o un cuadrado:

Código HTML	Resultado
<pre>&lt;UL type = disk &gt; &lt;LI&gt;España &lt;LI&gt;Francia &lt;LI&gt;Italia &lt;LI&gt;Portugal &lt;/UL&gt;</pre>	<ul style="list-style-type: none"> <li>• España</li> <li>• Francia</li> <li>• Italia</li> <li>• Portugal</li> </ul>
<pre>&lt;UL type = square&gt; &lt;LI&gt;España &lt;LI&gt;Francia &lt;LI&gt;Italia &lt;LI&gt;Portugal &lt;/UL&gt;</pre>	<ul style="list-style-type: none"> <li><input type="checkbox"/> España</li> <li><input type="checkbox"/> Francia</li> <li><input type="checkbox"/> Italia</li> <li><input type="checkbox"/> Portugal</li> </ul>

Las listas de definición muestran los elementos como en un diccionario, o sea, el término y la definición. Se utiliza para ellas la etiqueta `<DL></DL>`. El elemento marcado como término se antecede de la etiqueta `<DT>`, el marcado como definición se antepone de la etiqueta `<DD>`.

Código HTML	Resultado
<pre>&lt;DL&gt; &lt;DT&gt;WWW &lt;DD&gt;Abreviatura de World Wide Web &lt;DT&gt;FTP &lt;DD&gt;Abreviatura de File Transfer Protocol &lt;DT&gt;IRC &lt;DD&gt;Abreviatura de Internet Relay Chat &lt;/DL&gt;</pre>	<pre>WWW Abreviatura de World Wide Web FTP Abreviatura de File Transfer Protocol IRC Abreviatura de Internet Relay Chat</pre>

Existen dos tipos de listas menos comunes. Las listas de Menú o Directorio se comportan igual que las listas sin numerar. La lista de Menú utiliza la etiqueta `<MENU></MENU>` y los elementos se anteceden de `<LI>`. El resultado es una lista sin numerar mas "compacta" es decir, con menos espacios interlineales entre los elementos. La lista de Directorio utiliza la etiqueta `<DIR></DIR>` y los elementos se anteceden de `<LI>`. Los elementos tienen un límite de 20 caracteres. Todas las listas se pueden "anidar", es decir incluir una lista dentro de otra, con lo que se consigue una estructura como la siguiente:

Código HTML	Resultado
<pre>&lt;UL type= disk&gt; &lt;LI&gt;Buscadores   &lt;UL&gt;     &lt;LI&gt;Yahoo     &lt;LI&gt;Ole     &lt;LI&gt;Lycos   &lt;/UL&gt; &lt;LI&gt;Ligas   &lt;UL&gt;     &lt;LI&gt;Microsoft     &lt;LI&gt;IBM   &lt;/UL&gt; &lt;/UL&gt;</pre>	<pre>• Buscadores   ☐ Yahoo   ☐ Ole   ☐ Lycos • Ligas   ☐ Microsoft   ☐ IBM</pre>

### 3.5.7. Imágenes

Para incluir una imagen en el código HTML, se emplea la etiqueta `<IMG>`. Existen dos formatos de imágenes que todos los visualizadores modernos reconocen: **GIF** y **JPG**. Cualquier otro tipo de archivo gráfico o de imagen (BMP, PCX, CDR, etc...) no será mostrado por el visualizador, a no ser que se tenga un programa externo que permita su despliegue.

La etiqueta **<IMG>** tiene varios atributos:

- **src** = "imagen"  
Indica el nombre del archivo gráfico a mostrar.
- **alt** = "Texto"  
Muestra el texto indicado en caso de que el visualizador a emplear no sea capaz de visualizar la imagen.
- **align** = TOP / MIDDLE / BOTTOM  
Indica como se alinea el texto que sigue a la imagen. TOP alinea el texto con la parte superior de la imagen, MIDDLE con la parte central, y BOTTOM con la parte inferior.
- **border** = tamaño  
Indica el tamaño del "borde" de la imagen. A toda imagen se le asigna un borde que es visible cuando la imagen forme parte de un Hyperenlace.
- **height** = tamaño  
Indica el alto de la imagen en puntos o en porcentaje. Se usa para variar el tamaño de la imagen original.
- **width** = tamaño  
Indica el ancho de la imagen en puntos o en porcentaje. Se usa para variar el tamaño de la imagen original.
- **hspace** = margen  
Indica el número de espacios horizontales, en puntos, que separan la imagen del texto que le siga y anteceda.
- **vspace** = margen  
Indica el número de puntos verticales que separaran la imagen del texto que le siga y anteceda.

Por ejemplo:

Código HTML	Resultado
<code>&lt;IMG src="saludo.gif" alt= "Hola !!" &gt;</code>	

La imagen a mostrar puede encontrarse en el mismo lugar (URL) que la página Web; si este no fuera el caso, el nombre de la imagen debe contener el URL donde se encuentre la imagen.

Ejemplo:

```
<IMG src="http://www.microsoft.com/iexplorer.gif">
```

Varios ejemplos con tamaños de imagen y alineación, se muestran a continuación. (Recuerde que en función del visualizador empleado, pueden verse o no los efectos de cada atributo).

Código HTML	Resultado
<code>&lt;IMG src="saludo.gif" align= BOTTOM&gt;</code> Hola !!!	 Hola !!!
<code>&lt;IMG src="saludo.gif" height=20 &gt;</code>	
<code>Bienvenido&lt;IMG src="saludo.gif" hspace=20&gt;a este espacio.</code>	Bienvenido  a este espacio.

### 3.5.8. Hyperenlaces.

La característica principal de una página Web es que podemos incluir Hyperenlaces. Un Hyperenlace es un elemento de la página que hace que el visualizador acceda a otro recurso, otra página Web, un archivo, o un sitio, etc...

Para incluir un Hyperenlace se utiliza la etiqueta `<A></A>`. El texto o imagen que se encuentre dentro de los límites de esta etiqueta será sensible, esto quiere decir que al presionar con el ratón sobre él, se realizará la función de hyperenlace indicada por la etiqueta `<A></A>`. Si el Hyperenlace está indicado para un texto, este aparecerá subrayado y en distinto color, si se trata de una imagen, esta aparecerá con un borde rodeándola a menos que se especifique lo contrario. Esta etiqueta tiene el atributo `href` que indica el lugar a donde llevará el Hyperenlace al presionar con el ratón. Por ejemplo:

`<A href = "http://morsa.dgscs.unam.mx/"> Presione aquí para ir a la página de Productos Interactivos para la Docencia en la DGSCA</A>`

Resultado:

[Presione aquí para ir a la página de Productos Interactivos para la Docencia en la DGSCA](http://morsa.dgscs.unam.mx/)

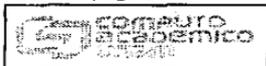
Al colocar el ratón sobre la frase y presionar, el visualizador accederá a la página Web indicada por el atributo `href`, es decir, accede a la página situada en el servidor `http://morsa.dgscs.unam.mx/`

El mismo efecto se puede producir para un gráfico, por ejemplo para acceder a la página informativa de la DGSCA: Presione la imagen para entrar a la página informativa de

Dirección General de Servicios de Cómputo Académico <A href = "http://www.museovirtual.unam.mx/dgsca.html" ><IMG src = "dgsca.gif"></A>

Lo cual da como resultado:

Presione la imagen para entrar a la página informativa de Dirección General de Servicios



de Cómputo Académico

Un Hyperenlace también puede llevar a una zona de una página. Para ello se debe marcar en la página las diferentes secciones en las que se divide, lo cual se hace con el atributo **name**. Por ejemplo:

```
<A name = "seccion1"></A>
```

Esta instrucción marca el inicio de una sección dentro de una página. La sección se llamará seccion1.

Para lo cual se emplea: <A href = "#seccion1">Primera Parte</A>

O también:

```
<A href = "http://morsa.dgsca.unam.mx/mipagina.htm#seccion1">Primera Parte</A>
```

Un Hyperenlace puede hacerse a cualquier tipo de archivo. Con las etiquetas anteriores se ha visto como hacer enlaces a páginas Web o secciones dentro de una; pero también se puede hacer un Hyperenlace a un grupo de noticias, o a cualquier otro servicio de Internet, siempre que sea soportado por el visualizador que estemos empleando. Por ejemplo:

```
<A href = "ftp://ftp.twcowns.com/">sitio de shareware en WWW</A>
```

Asimismo podemos hacer que el Hyperenlace de como resultado el envío de un correo electrónico a una dirección de correo determinada. Por ejemplo:

```
<A href = "mailto:ulises@servidor.unam.mx">Envíame tus sugerencias</A>
```

También podemos realizar un Hyperenlace a un archivo cualquiera; en este caso el visualizador intentará "ejecutar" el archivo, y si este NO puede leerlo, preguntará si deseamos copiarlo a nuestra computadora, en donde se ejecuta el programa visualizador:

```
<A href = "manual.zip">Presione aquí para obtener una copia del manual</A>
```

### 3.5.9. Tablas

Hasta que no se empezaron a usar las tablas, la única manera de tabular las cosas era utilizar la etiqueta de preformateado, con la que es necesario colocar manualmente los espacios en blanco para que quede todo alineado formando filas y columnas, con un resultado muy poco estético.

#### **Estructura de una tabla:**

Las etiquetas necesarias para crear una tabla son:

- La etiqueta general, que engloba a todas las demás es `<TABLE>` y `</TABLE>`. Es decir:

```
<TABLE>
[resto de las etiquetas]
</TABLE>
```

Con esto se presentan los datos tabulados, pero faltaría la característica que hace más atractiva a la tabla, y es que los datos vayan dentro de espacios delimitados por un borde; para esto se tiene que añadir el atributo `BORDER` a la etiqueta, es decir:

```
<TABLE BORDER>
[resto de las etiquetas]
</TABLE>
```

- En el siguiente nivel, dentro del anterior, se encuentran las etiquetas para formar cada fila (row) de la tabla, que son `<TR>` y `</TR>`. Es necesario repetir tantas veces como filas queremos que tenga la tabla; es decir, para una tabla con dos filas:

```
<TR>
[etiquetas de las distintas celdas de la primera fila]
</TR>
<TR>
[etiquetas de las distintas celdas de la segunda fila]
</TR>
```

- En el último nivel (dentro de las anteriores), se encuentran las etiquetas de cada celda, que son `<TD>` y `</TD>`, las cuales engloban el contenido de cada celda concreta (texto, imágenes, etc.).

Un ejemplo de una tabla con dos filas (cada fila tiene tres celdas).

```
<TABLE BORDER>
<TR>
<TD>fila1-celda1</TD> <TD>fila1-celda2</TD> <TD>fila1-celda3</TD>
</TR>
```

```
<TR>
<TD>fila2-celda1</TD> <TD>fila2-celda2</TD> <TD>fila2-celda3</TD>
</TR>
</TABLE>
```

Esto resulta en:

fila1-celda1	fila1-celda2	fila1-celda3
fila2-celda1	fila2-celda2	fila2-celda3

### **Filas con desigual número de celdas:**

Para el siguiente ejemplo se han colocado dos filas con igual número de celdas, si este número es distinto, el visualizador forma el número de filas y columnas que haga falta, dejando espacios en blanco en las filas que tengan menos celdas. Si en el ejemplo anterior se eliminó la tercera celda de la segunda fila, (si borramos `<TD>fila2-celda3</TD>`), entonces se obtiene:

fila1-celda1	fila1-celda2	fila1-celda3
fila2-celda1	fila2-celda2	

### **Título de la tabla:**

Se puede añadir un título a la tabla, es decir un texto situado encima de la tabla que indique cuál es su contenido, lo cual se consigue con las etiquetas `<CAPTION>` y `</CAPTION>`.

### **Variando el tamaño de los bordes:**

El atributo `BORDER` coloca por omisión un borde de espesor igual a la unidad, pero se puede hacer que este borde sea tan grueso como uno quiera:

```
<TABLE BORDER=número deseado>
```

Si en el ejemplo anterior se coloca:

```
<TABLE BORDER=5>
```

Entonces:

fila1-celda1	fila1-celda2	fila1-celda3
fila2-celda1	fila2-celda2	

### **Celdas encabezado:**

Además de las celdas que contienen datos normales, se pueden añadir, celdas de encabezado (header), que se distinguen por texto resaltado y centrado. Esto se consigue

### Creación de Páginas en WWW

con la etiqueta <TH> y </TH> (en vez de la normal <TD> y </TD>). Sustituyendo en el ejemplo anterior, una fila de estas celdas de encabezado:

```
<TR>
<TH>Columna 1</TH> <TH>Columna 2</TH> <TH>Columna 3</TH>
</TR>
```

Se obtiene (en el ejemplo de filas desiguales):

Columna 1	Columna 2	Columna 3
fila1-celda1	fila1-celda2	fila1-celda3
fila2-celda1	fila2-celda2	

### Contenido de las celdas:

Dentro de estas se puede colocar cualquier elemento HTML, como imágenes, enlaces, etc., para esto, basta con colocar dentro de la etiqueta de la celda la correspondiente a una imagen, un enlace, u otro elemento.

Con una imagen:

```
<TABLE BORDER=2>
<TR><TD><IMG SRC="babe.gif"></TD></TR>
</TABLE>
```

Con un enlace:

```
<TABLE BORDER=2>
<TR><TD><A HREF="index.html">Página principal </A></TD></TR>
</TABLE>
```

### Posición del contenido dentro de una celda:

Normalmente el contenido de una celda se encuentra alineado a la izquierda, sin embargo se puede cambiar esto añadiendo dentro de la etiqueta de la celda los siguientes atributos:

```
<TD ALIGN=CENTER> Al centro </TD>
<TD ALIGN=RIGHT> A la derecha </TD>
<TH ALIGN=LEFT> Encabezado a la izquierda </TH>
```

La alineación por omisión en sentido vertical es en medio, sin embargo es modificable, añadiendo dentro de la etiqueta de la celda los siguientes atributos:

```
<TD VALIGN=TOP> Arriba </TD>
<TD VALIGN=BOTTOM> Abajo </TD>
```

	Arriba	
		Abajo

**Dimensiones de una tabla:**

El visualizador se encarga normalmente del dimensionado en el tamaño total de una tabla de acuerdo con el número de filas y columnas, por el contenido de las celdas, espesor de los bordes, etc. Sin embargo en ocasiones conviene modificar estos valores para que la tabla tenga ciertas dimensiones personalizadas según convenga, esto se consigue añadiendo dentro de la etiqueta de la tabla los atributos **WIDTH** y **HEIGHT** igual a un porcentaje de la dimensión de la pantalla, o a una cifra que equivale al número de pixels. Por ejemplo, si en el último ejemplo colocamos:

```
<TABLE WIDTH=60%>
```

Obtenemos:

	Arriba	
		Abajo

O, bien:

```
<TABLE HEIGHT=200>
```

Con lo cual se obtiene:

	Arriba	
		Abajo

**Celdas que abarcan a otras:**

Esto se consigue añadiendo dentro de la etiqueta de la celda los atributos **COLSPAN=numero** para extenderse sobre un número determinado de columnas, o **ROWSPAN=numero** para extenderse verticalmente sobre un número determinado de filas. Para el primer ejemplo de tabla descrito en esta sección, vamos a añadir una fila con una sola celda la cual abarca a dos columnas, de la forma siguiente:

```
<TR> <TD COLSPAN=2> Celda sobre 2 columnas </TD> <TR>
```

*Creación de Páginas en WWW*

Celda sobre 2 columnas		
fila1-celda1	fila1-celda2	fila1-celda3
fila2-celda1	fila2-celda2	fila2-celda3

O bien en la misma tabla, para añadir una celda en la primera fila, pero que abarque también a la siguiente se debe implementar:

`<TD ROWSPAN=2> Celda junto a 2 filas </TD>`

# Capítulo IV

## Extensiones HTML

---

### 4.1. Formas

Las formas permiten dentro de una página Web solicitar información al usuario y procesarla. En una forma podremos solicitar diferentes tipos de datos (campos) cada uno de los cuales queda asociado a una variable, una vez introducidos los valores en los campos el contenido de estos será enviado a la dirección (URL), donde se tiene el programa que procesar el valor de las variables, para realizar este último paso, es necesario crear un programa en algún lenguaje de programación como PERL, C, Visual Basic o C++, entre otros. A este programa externo se llama CGI (Common Gateway Interface).

La declaración de la forma se coloca entre las etiquetas `<FORM>` `</FORM>`. En la declaración se indican las variables de entrada. La etiqueta `<FORM>` tiene los atributos *action* y *method*.

**action** = "programa"

Indica el programa que va a "procesar" las variables que se envíen con la forma, por ejemplo para el caso de un correo electrónico será:

`action="mailto:ulises@servidor.unam.mx"`

**method** = POST / GET

Indica el método por el cual se transferirán las variables. **POST** produce la modificación del documento destino (como en el caso de enviar por correo las variables). **GET** no produce cambios en el documento destino (por ejemplo, una consulta a una base de datos, en algún visualizador como Yahoo o Lycos).

## Campos de entrada:

Para introducir las variables se utiliza la etiqueta **<INPUT>**. Esta etiqueta tiene el atributo **type** que indica el tipo de variable a introducir y **name** que indica el nombre que se le dará al campo. Cada tipo de variable tiene sus propios atributos.

- **type = text name = campo**  
Indica que el campo a introducir será un texto. Sus atributos son :
  - maxlength = numero**  
Número máximo de caracteres a introducir en el campo.
  - size = numero**  
Tamaño en caracteres que se muestra en pantalla.
  - value = "texto"**  
Valor inicial del campo. Normalmente será " " (o sea, nulo).
- **type = password name = campo**  
Indica que el campo será una palabra que NO se verá en el visualizador, muestra asteriscos (\*) en lugar de letras. Sus atributos opcionales son los mismos que para text.
- **type = checkbox name = campo**  
El campo se elegirá marcando una casilla (permitiéndose marcar varias casillas). Los valores de las casillas serán indicados por :
  - value = "valor"**
  - checked**  
La casilla aparecerá marcada por omisión.
- **type = radio name = campo**  
El campo se elegirá marcando una casilla (permite marcar una sola de las casillas). Los valores de las casillas serán indicados por :
  - value = "valor"**
- **type = image name = campo**  
El campo contendrá el valor de las coordenadas del punto de la imagen seleccionado, debe indicarse la imagen con el atributo:
  - src = "archivo de imagen".**
- **type = hidden name = campo**  
El usuario no puede modificar su valor, ya que el campo no es visible, este se manda siempre con el valor indicado por el atributo : **value = "valor"**

- **type = submit**

Representa un botón, al presionarlo, la información de todos los campos en la forma, se envía al programa indicado en **<FORM>**. Tiene el atributo **value = "texto"**, el cual indica el texto que aparecerá en el botón.

- **type = reset**

Representa un botón, al presionarlo, se borra el contenido de todos los campos contenidos en la forma. El atributo **value = "texto"** indica el texto que aparecerá en el botón.

### **Campos de selección:**

Este tipo de campos muestran una lista de opciones, entre las que debemos escoger una o varias. Para ello se emplean las etiquetas **<SELECT>** **</SELECT>**. Sus atributos son :

- **name = campo**

Nombre del campo

- **size = numero**

Numero de opciones visibles. Si se coloca 1 se presenta como un menú desplegable, al indicar más de uno se presenta como una lista con barra de desplazamiento.

- **multiple**

Permite seleccionar más de un valor para el campo.

Las diferentes opciones de la lista se indican con la etiqueta **<OPTION>**. Esta puede incluir el atributo *selected* para indicar cual es la opción por omisión, en caso de que no se especifique, se toma por omisión la primera opción de la lista.

### **Areas de texto:**

Representa un campo de texto de múltiples líneas, normalmente empleada para incluir en él comentarios, la etiqueta usada es **<TEXTAREA>** **</TEXTAREA>**, y sus atributos :

- **name = campo**

Nombre del campo.

- **cols = numero**

Numero de columnas de texto visibles.

- **rows = número**  
Número de filas de texto visibles.
- **wrap = VIRTUAL / PHYSICAL**  
Justifica el texto automáticamente en el interior de la caja (área de texto). La opción PHYSICAL envía las líneas de texto separadas en líneas físicas. La opción VIRTUAL envía todo el texto continuo.

*A continuación presento un ejemplo completo de forma utilizando varios elementos de introducción de datos:*

```
<FORM action = "mailto: ulises@servidor.unam.mx" method = post >
Su Nombre:<INPUT type = text name = nombre size = 30 >
Su Clave: <INPUT type = password name = clave size = 8 >
<P>
Archivos a Enviar:
<INPUT type = checkbox name = archivo value = "JavaScript" > Introducción a JavaScript
<INPUT type = checkbox name = archivo value = "curso_c" > Curso de C
<INPUT type = checkbox name = archivo value = "curso_unix" > Curso de Unix
<P>
Su Edad :
<INPUT type = radio name = edad value = "-18" > Menos de 18 años
<INPUT type = radio name = edad value = "18-35" > Entre 18 y 35 años
<INPUT type = radio name = edad value = "+35" > Mas de 35 años
<P>
<INPUT type = hidden name = lugar value = "home" >
Como encontraste este sitio:
<SELECT name = donde >
<OPTION>Gaceta UNAM
<OPTION>Por el buscador Yahoo
<OPTION>Recomendación
<OPTION>Por casualidad
</SELECT>
<P>
Por favor tus Comentarios:
<BR>
<TEXTAREA name = comentario rows = 5 cols = 40 wrap = virtual ></TEXTAREA>
<P>
<INPUT type = submit value = "Enviar" >
<INPUT type = reset value = "Borrar" >
</FORM>
```

*El resultado directamente a través del visualizador se ve parecido al de la siguiente figura:*

**Ejemplo de forma, interpretada por un visualizador.**

*El texto que se recibe por correo electrónico es parecido a:*

```
nombre=Pedro+Perez&clave=12345678&archivo=JavaScript&archivo=curso_c&edad=18-35&lugar=pagina+home&donde=Gaceta+UNAM&comentario%04=Espero+que+por+favor+mande+los%0D%0Aarchivos+antes+del+martes%0D%0A%0D%0ASaludos.%0D%0A
```

Se puede observar que en el caso de correo electrónico, se separan las variables con el símbolo **&**, los espacios se sustituyen por el signo **+** y se representan los códigos de retorno de carro y avance de línea del campo de texto con los caracteres **%0D** y **%0A** respectivamente.

Si en vez de enviar estas variables por correo electrónico, fuesen enviadas a un programa (CGI) en el servidor, este puede procesarlas e interactuar directamente con el cliente (o visualizador).

## 4.2. Common Gateway Interface (CGI)

Un CGI es un programa que se ejecuta directamente en el servidor Web, sirven para procesar información a través de variables propias de este tipo de programas, además se pueden generar documentos html en forma automática, consulta a bases de datos

remotas, a bancos de datos distribuidos por todo Internet, pero principalmente permiten una interacción en tiempo real por parte del cliente hacia el servidor.

Una de las principales utilidades del CGI es procesar los resultados de una forma (*FORM*). Un CGI puede ser escritos en cualquier lenguaje de programación y son ejecutados en el servidor cuando se efectúa una referencia a este. El programa debe ser ejecutable en el servidor (para el caso del S.O. UNIX), es decir debe tener los permisos necesarios para ejecutar el o los programas residentes. Los lenguajes más comunes para escribir estos programas son PERL, C ó bien *scripts* de cualquier shell (para el caso de UNIX: *csh, sh, ksh* ó *zsh*).

El programa CGI debe encontrarse en un directorio especial llamado *cgi-bin*, destinado a contener el o los compiladores apropiados para hacer las llamadas necesarias, este directorio es asignado por omisión en el momento de instalar el software de servidor Web.

### **Variables de entorno:**

Para pasar datos del servidor al CGI, se emplea una serie de variables de entorno, las cuales quedan definidas al ejecutar el script. Las variables de entorno comúnmente empleadas son:

- **REQUEST\_METHOD**

Es el método por el cual se realiza la llamada al script, este puede ser GET ó POST como se explicó anteriormente al hablar de formas.

- **PATH\_INFO**

El CGI puede ser llamado directamente desde el cliente, en esta variable se debe incluir la ruta, por ejemplo:

```
http://www2.uca.mx/binarios/cgi-bin.cgi/ruta/archivo
```

```
PATH_INFO = /ruta/archivo
```

La información que se incluye será pasada de forma codificada al CGI, de forma que los espacios se convierten en signos '+' y los caracteres especiales se codifican de la forma %xx, donde xx el código ASCII en hexadecimal del caracter. Esto debe ser tomado en cuenta cuando se interprete la información obtenida en formato URL.

- **PATH\_TRANSLATED**

Una versión decodificada de PATH\_INFO, la cual obtiene la ruta y la convierte de ruta virtual (en referencia del servidor) a ruta real.

- **SCRIPT\_NAME**  
Nombre por el que fue llamado el CGI.
- **QUERY\_STRING**  
La información que sigue al símbolo ? en el URL que hace referencia al CGI. No necesita ser decodificada de ninguna manera.  
  
Cuando se utiliza el método GET las variables de la forma se pueden interpretar con esta variable de entorno.  
  
El formato en que envían estas variables es el siguiente:  
  
*?variable1=valor1&variable2=valor2& ... &variablen=valorn*
- **REMOTE\_HOST**  
Es la computadora desde la que se ejecuta el CGI.
- **REMOTE\_ADDR**  
Dirección IP de la computadora remota que ejecuto el CGI.
- **CONTENT\_LENGTH**  
El número de caracteres enviados por el cliente al CGI.
- **HTTP\_USER\_AGENT**  
Cliente WWW que ejecuta el CGI, el formato general es:  
  
*programa/versión librería/versión*  
  
El CGI podría enviar un distinto código HTML según cada cliente.

### **Entrada de datos al CGI:**

Para las formas que utilizan el método **POST**, la información es enviada al CGI por la entrada estándar (*stdin*). El servidor envía las variables de entorno **CONTENT\_LENGTH** y **CONTENT\_TYPE** indicando la longitud y el tipo de datos enviados. El servidor no está obligado a enviar una marca de final de archivo tras los datos enviados, por tanto la variable **CONTENT\_LENGTH** debe ser tomada en cuenta para determinar la cantidad de datos enviados. **CONTENT\_TYPE** suele contener el valor:

*application/x-www-form-urlencoded*

que indica que se trata de información codificada del URL.

## Salida de datos del CGI:

El CGI debe enviar los datos a la salida estándar (*stdout*), esta puede ser un documento generado por el CGI o instrucciones al servidor para obtener el documento correcto. La salida del CGI debe comenzar con un pequeño encabezado que identifica al documento. Las etiquetas que definen este encabezado son:

### Content-type:

Es el tipo MIME del documento que se devuelve. Si toma el valor *text/html*, la salida se interpreta como código html; por tanto la salida del CGI puede contener instrucciones HTML que son interpretadas correctamente por el cliente. Otros valores que puede tomar son: *text/plain* para texto normal ó *image/gif*, en donde la salida se interpreta como una imagen GIF.

Por ejemplo:

```
printf ("Content-type: text/html\n");
```

### Location:

Indica al servidor que se esta enviando la referencia a un documento en lugar del documento en sí. Si el valor es un URL, el servidor indica al cliente que debe realizar un direccionamiento a esa dirección.

Esta permitido el uso de información adicional (*PATH\_INFO*) y la variable *?*, pero no la variable *#*, para indicar un punto determinado dentro de un documento.

### Status:

Indica un código de estado para indicar errores, el formato es *nnn xxxxxxx*, donde *nnn* es un número de tres dígitos y *xxxxxxx* es una cadena de caracteres. Tras el encabezado se debe incluir dos caracteres de retorno de carro (*\n*), condición obligatoria para que se interprete apropiadamente el encabezado CGI, si no se respeta esto la salida del CGI podría generar el error: "Server Error 500".

Después de definir el encabezado se puede generar la salida del CGI, la cual se interpreta según el valor dado a *Content-type*.

### 4.3. Frames

Con los frames es posible dividir la ventana del visualizador en varias áreas, permitiendo mostrar un URL diferente en cada una de ellas. Cada frame permite:

- Mostrar un propio URL, diferenciado del resto de los frames de la pantalla, de esta manera una hiperliga puede tener como destino un documento.
- Tiene asociado un nombre, que los distingue del resto de los frames del total en pantalla y permite el uso de hiperligas.
- Se puede modificar el tamaño de la ventana, así el frame queda ajustable.

Estas características permiten crear nuevos tipos de documentos, en los que por ejemplo, se mantiene una región estática (listas de enlaces, barra de botones, formas) y otra zona dinámica en la que se muestra el contenido del URL. Una de las principales utilidades de los frames es la creación de páginas con un índice o un encabezado estático, consiguiendo así una navegación con el poder de acceder al índice de manera más rápida y efectiva.

El uso de los frames es útil para cierto tipo de documentos, sin embargo llega a dificultar la visualización, es decir, dentro de un documento con frames no se tiene habilitados los botones de documento previo (*back*), y documento siguiente (*forward*), ya que ambos trasladan a la ventana activa fuera del documento con frames. Para ver el documento previo dentro de un frame, se debe utilizar el botón derecho del ratón y seleccionar la opción: *Back in Frame*.

No todos los visualizadores pueden mostrar documentos con frames, sólo son interpretados por Navigator 2.0 o superior e Internet Explorer 3.0. En este caso si se desea que el documento sea accesible para una gran cantidad de usuarios en el Web, se debe crear un documento con frames y otro alternativo sin estos, empleando la etiqueta *NOFRAMES*. Un documento con frames tiene una estructura del documento distinta, NO se define la etiqueta *BODY*. Su estructura es la siguiente:

```
<HTML>
  <HEAD>
    Definiciones del encabezado
  </HEAD>
    <FRAMESET>
      Definición de los frames que forman el documento y de los
      archivo que incluye cada una.
    </FRAMESET>
    <NOFRAMES>
      Instrucciones HTML que se muestra en los visualizadores que
      no soporten frames.
    </NOFRAMES>
</HTML>
```

Dentro de la etiqueta NOFRAMES se puede incluir un mensaje, explicando que el documento sólo se visualiza con navegadores que soporten frames.

#### <FRAMESET ...> definición.

Con esta instrucción se definen los frames que formarán el documento, su sintaxis es similar a las tablas, Su formato es el siguiente:

```
<FRAMESET ROWS=Lista de las Filas COLS=Lista de las Columnas>
  <FRAME SRC=URL_1 NAME="Nombre de el frame1">
  <FRAME SRC=URL_2 NAME="Nombre de el frame2">
  ...
  <FRAME SRC=URL_N NAME="Nombre de el frameN">
</FRAMESET>
```

Se define sólo uno de los atributos: renglones (**ROWS**), o columnas (**COLS**).

#### **ROWS:**

Se define separado por comas el tamaño de cada una de los frames; así se divide la pantalla en forma horizontal (según cada una de las filas definidas). El tamaño del frame, puede expresarse de las siguientes formas:

- En *valor absoluto*, que indicará el tamaño en puntos de la pantalla. En este caso si todas los frames se indican de este modo, los valores se ajustan automáticamente para que los frames ocupen la totalidad del espacio de la ventana del visualizador.
- En *tanto por ciento* sobre el tamaño de la ventana, en este caso si los porcentajes suman un valor distinto del 100%, se ajustarán para que coincidan con el tamaño de la ventana.
- De forma *relativa* con el símbolo \* que indica el tamaño restante de la ventana. Por ejemplo, al indicar una frame como 20% y otro con \*, el último ocupará un 80% de espacio restante.

Observe el siguiente ejemplo:

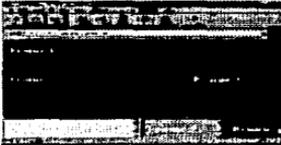
Código HTML	Resultado
<pre>&lt;FRAMESET ROWS="100, 50%, *&gt; &lt;FRAME SRC="frame1.htm" NAME="frame1"&gt; &lt;FRAME SRC="frame2.htm" NAME="frame2"&gt; &lt;FRAME SRC="frame3.htm" NAME="frame3"&gt; &lt;/FRAMESET&gt;</pre>	
<p>Se definen tres frames horizontales: el primero ocupa 100 puntos de la pantalla, el segundo 50% del tamaño de la ventana del visualizador y el tercero el espacio restante, si se varía el tamaño de la ventana del visualizador cambia el tamaño de las dos últimas.</p>	

**COLS:**

Toma los mismos valores que ROWS, se definen los frames en forma vertical.

Código HTML	Resultado
<pre>&lt;FRAMESET COLS="100, 50%, *&gt; &lt;FRAME SRC="frame1.htm" NAME="frame1"&gt; &lt;FRAME SRC="frame2.htm" NAME="frame2"&gt; &lt;FRAME SRC="frame3.htm" NAME="frame3"&gt; &lt;/FRAMESET&gt;</pre>	
<p>Se definen los frames de forma vertical, guardando la misma proporción que en el caso anterior.</p>	

Una vez definido el número de frames por fila o por columna se define el contenido de cada uno de estos con la etiqueta **FRAME**, para anidar frames se emplea **FRAMESET**, la cual divide en el número de frames indicados, como se ve en el siguiente ejemplo:

Código HTML	Resultado
<pre>&lt;FRAMESET ROWS="100, 50%, *&gt; &lt;FRAME SRC="frame1.htm" NAME="frame1"&gt; &lt;FRAMESET COLS="2", *&gt;   &lt;FRAME SRC="frame2.htm" NAME="frame2"&gt;   &lt;FRAME SRC="frame3.htm" NAME="frame3"&gt; &lt;/FRAMESET&gt; &lt;FRAMESET COLS="50%, 30%, 20%*&gt;   &lt;FRAME SRC="frame4.htm" NAME="frame4"&gt;   &lt;FRAME SRC="frame5.htm" NAME="frame5"&gt;   &lt;FRAME SRC="frame6.htm" NAME="frame6"&gt; &lt;/FRAMESET&gt; &lt;/FRAMESET&gt;</pre>	
<p>Tomando como base los frames del ejemplo anterior, se dividen dos de los frames, formando en este caso un documento con 6 divisiones, cada uno muestra un archivo distinto y tienen un nombre distinto.</p>	

#### **<FRAME ...> definición de frame**

La etiqueta FRAME que define el documento presenta además algunos otros atributos como:

```
<FRAME SRC="URL" NAME="Nombre de el frame" MARGINWIDTH="ancho del margen"
MARGINHEIGHT="alto del margen" SCROLLING=YES | NO | AUTO NORESIZE>
```

- **URL:**  
Especifica el documento HTML o archivo que se muestra en el frame definido. Si no se especifica documento alguno se muestra como vacío.
- **NAME:**  
Indica el nombre del frame, este nombre es importante ya que se usa en las hiperligas (normalmente en los documentos de los otros frames) para indicar el destino del documento, al no indicar el nombre sólo se podrá mostrar el documento actual, sin que sea posible cambiarlo mediante hiperligas.
- **MARGINWIDTH:**  
Indica el ancho del margen, este atributo es opcional y normalmente el visualizador ajusta todos los márgenes al mismo tamaño, si se especifica su valor será en puntos de la pantalla.
- **MARGINHEIGHT:**  
Igual que en el caso anterior pero para el alto de los márgenes. Lo normal es no especificar ninguno de estos dos atributos.

- SCROLLING:**  
 Indica si el frame tendrá o no una barra de desplazamiento, esta barra se muestra lateralmente y permite desplazarse por el documento presionando con el ratón en ella, al tomar el valor YES siempre se muestra, para el valor AUTO sólo se muestra si el documento no cabe en el frame. Y por último el valor NO indica que en ningún caso se muestre la barra de scroll. Al no indicar nada se toma por omisión el valor AUTO.
- NORESIZE:**  
 Indica que el frame no puede ser variado de tamaño por el usuario, se puede variar el tamaño del frame situando el cursor del ratón por encima y arrastrándolo en la dirección deseada. Con este atributo el usuario no puede cambiar el tamaño de el frame en ningún caso. Por omisión todos los frames pueden variar su tamaño.

### **TARGET: frame destino**

El uso de frames introduce un nuevo atributo a alguna de las etiquetas especificadas con anterioridad, este atributo es **TARGET**, el cual indicará el frame destino de la operación. Normalmente en páginas sin estos, cuando se siguen hiperligas, este se muestra en la ventana del visualizador sustituyendo el documento actual. Con los frames se puede especificar cual es el destino. Para especificar el nombre del frame se emplea el atributo **NAME**. Las etiquetas que permiten el uso de **TARGET** son las siguientes:

- A**  
 En las hiperligas indicará el frame donde se muestra el documento, una vez que se siga la hiperliga.  

```
<A HREF="url" TARGET="frame">
```
- BASE:**  
 Indicará el frame en la que se muestra por omisión todos las hiperligas del documento actual. Se debe especificar en el encabezado del documento (HEAD).  

```
<BASE TARGET="frame">
```
- AREA:**  
 En la definición de imágenes sensibles en el cliente, se indica el frame donde se verá el documento que se activa en la zona correspondiente de la imagen.  

```
<AREA SHAPE=RECT COORDS="x,y,..." HREF="url" TARGET="frame">
```
- FORM:**  
 Indicará el frame destino del resultado en la forma.  

```
<FORM ACTION="url" TARGET="frame">
```

Existen ciertos valores especiales de TARGET, que permiten definir valores distintos a los que los frames definen:

- TARGET="\_blank": Indica que se muestre una nueva ventana vacía.
- TARGET="\_self": Muestra en la misma ventana o frame que hace la llamada, se puede usar para modificar el valor dado por BASE.
- TARGET="\_parent": Muestra en el frame o estructura de frames que llamó al documento actual.
- TARGET="\_top": Indica que se muestre la ventana completa, eliminando la estructura de frames que tenga esta.

## 4.4. Etiquetas Especiales

Netscape y Microsoft han añadido a sus estándares de HTML (en sus navegadores), diversas etiquetas para hacer más atractiva la visualización de las páginas WWW. Estas etiquetas pueden no funcionar en algunos visualizadores específicos, sin embargo la implementación y uso de ellas por parte de estas compañías, líderes en el mercado de software para Internet, hace prever que serán incluidas en las nuevas revisiones que se hagan a los estándares de HTML (en sus niveles 3 y futuros), así como para el resto de los visualizadores clientes comunes.

### Applet

La etiqueta `<APPLET></APPLET>`, indica la ejecución de un programa (applet) externo escrito en lenguaje **JAVA**. Java es un lenguaje de programación creado por Sun Microsystems, que permite realizar operaciones multimedia sin la necesidad de emplear nuevas etiquetas HTML. Esta etiqueta tiene los siguientes atributos:

- **codebase = URL**  
Dirección donde se encuentra el "traductor" del applet Java ( Por ejemplo: `http://serpiente.dgsca.unam.mx/java`).
- **code = nombre\_programa**  
Indica el nombre del programa Java (applet) a ejecutar.

- **width** = num.
  - **height** = num.
- Indican el espacio (ancho y alto) en el que el programa realizará su función.

Dentro de la etiqueta **<APPLET>** se encuentra la etiqueta **<PARAM>** que envía al programa Java los parámetros necesarios para su evaluación. Esta etiqueta tiene los atributos :

- **name** = campo  
Nombre de la variable a enviar.
- **value** = valor  
Valor de la variable a enviar.

Un ejemplo en el que se ejecuta un programa Java que permite que un texto se desplace de un lado a otro en una zona de la pantalla es:

```
<APPLET codebase="http://serpiente.dgsca.unam.mx/java" code="Laufschrift.class" width
= 350 height = 25 >
<PARAM name = bg.color value = "0,255,0">
<PARAM name = message value = "***Bienvenido a este espacio WEB - UNAM -
DGSCA***"></APPLET>
```

## Marquee

La etiqueta **<MARQUEE></MARQUEE>**, crea un marco con un texto en su interior que se desplaza. Funciona únicamente con Explorer (de Microsoft). Sus atributos son los siguientes :

- **align** = top / middle / bottom  
Indica si el texto del interior de la marquesina se alinea en la zona alta (top), en la baja (bottom) o en el centro (middle) de la misma.
- **bgcolor** = "codigo de color"  
Indica el color del fondo del marco.
- **direction** = left / right  
Indica hacia que lugar se desplaza el texto, izquierda (left) o derecha (right).
- **height** = num o %  
Indica la altura del marco en puntos o porcentaje en función de la ventana del cliente.
- **width** = num o %  
Indica el ancho del marco en puntos o porcentaje en función de la ventana del visualizador.

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

- **loop** = num / infinite  
Indica el número de veces que se desplazará el texto por el marco. Si se indica infinite, se desplazará indefinidamente.
- **scrolldelay** = num.  
Indica el número de milisegundos que tarda en reescribirse el texto por el marco, a mayor número más lentamente se desplazará este.

Un ejemplo de esta etiqueta es:

```
<MARQUEE bgcolor = "#FFFFFF" width = 50% scrolldelay = 0 > Bienvenido a mi pagina personal</MARQUEE>
```

## Elementos multimedia

Las posibilidades multimedia del lenguaje HTML se limitan a insertar imágenes dentro de un texto, mostrar un fondo en el documentos y modificar los colores que presentan los distintos textos. Con las instrucciones de HTML que hemos visto hasta ahora, es lo máximo que se puede conseguir, **la posibilidad de mostrar video o sonido en un documento HTML esta limitada a la capacidad y a la configuración del cliente visualizador.**

La forma básica de incluir un archivo de un formato NO reconocido por el visualizador es incluirlo en una hiperliga, de forma que cuando el usuario siga la hiperliga el programa visualizador llame a una *aplicación externa* que trate este tipo de archivos. El visualizador tendrá definida una lista de aplicaciones que trabajen con los formatos más comunes de archivos, tanto de video, como audio e imágenes. Esta opción es poco funcional ya que no se pueden incluir en un documento común.

**Internet Explorer** de Microsoft incluye una serie de etiquetas y atributos que permiten el uso de elementos multimedia, sin embargo aún NO son estándar, y es imposible el despliegue o visualización en computadoras distintas a PC (configuración multimedia) con Windows.

**Navigator** de Netscape, proporciona extensiones a HTML con la inclusión de **Plug-ins** (aplicaciones creadas por terceros), estos son programas que se añaden al visualizador de Netscape versión 2.0 o superior y permiten manejar diversos tipos de archivos, de esta manera se pueden insertar estos tipos de archivos dentro de un documento HTML, permitiendo mostrar video en distintos formatos (mpeg, avi, mov), interpretar sonido, en varios formatos (au, mid), incluir archivos de Adobe Acrobat (pdf), mostrar archivos VRML (Virtual Reality Markup Language), etc. El principal problema es que estos plug-ins No son estándar ni están realizados por el mismo fabricante, además No todos son de libre distribución, y sólo son validos hasta ahora (durante la conclusión de este trabajo de tesis -invierno de 1996-), en PC's con Windows. Además la correcta visualización y despliegue del documento depende de que el usuario disponga del plug-in correspondiente instalado (apropiado según la aplicación). En la última versión de Navigator 3.0 (la liberación de

esta versión, se efectuó en el invierno de 1996), se incluyen por omisión, plug-in para audio, vídeo y VRML, por lo tanto cuando se extienda ampliamente el uso de esta versión, se podría asegurar la reproducción de este tipo de archivos.

## Extensiones multimedia de Internet Explorer

Internet Explorer de Microsoft permite mostrar vídeo en formato AVI (Vídeo for Windows) y reproducir sonido de una forma sencilla, para eso añade una nueva etiqueta y un atributo a la etiqueta de mostrar imágenes.

### <BGSOUND ...> reproducción de audio

<BGSOUND SRC="archivo de sonido" LOOP= n ó INFINITE>

Se incluye una sola vez esta etiqueta en el archivo que se desea tenga música de fondo, en el atributo **SRC** se indica el archivo de audio que se debe reproducir, inclusive el URL donde se localiza el archivo. Este puede ser en formato .wav, .au o MIDI (.mid). El atributo **LOOP** indica cuantas veces se ejecutará la pieza musical, puede indicarse un número o **INFINITE** de forma indefinida.

### <IMG DYN SRC ...> reproducción de vídeo

<IMG DYN SRC="archivo de sonido" LOOP= n ó INFINITE CONTROLS START= FILEOPEN ó MOUSEOVER>

Se trata de un nuevo atributo para la etiqueta **IMG** (en caso de que exista en el documento), en vez de mostrar una imagen, muestra un vídeo en formato AVI (Vídeo for Windows), se pueden emplear todos los atributos normales de **IMG**, inclusive usar las distintas alineaciones y formatos respecto al texto, además se puede variar el tamaño del vídeo con los atributos **HEIGHT** y **WIDTH**.

Se añaden otros atributos que sólo se pueden aplicar al vídeo, el atributo **IMG SRC**, indica el URL del archivo que se muestra, el atributo **LOOP**, indica el número de veces que se muestra el vídeo, el atributo **CONTROLS**, indica si se desean mostrar los botones de control, que permiten reproducir y controlar. El atributo **START** indica si la secuencia de vídeo empezará al abrir el archivo (**FILEOPEN**) o cuando se pase el cursor del ratón encima (**MOUSEOVER**).

De no poder mostrar el vídeo por tratarse de otro visualizador (por ejemplo Navigator), se puede indicar una imagen que se muestre de manera alternativa, incluyendo el atributo **SRC**.

## Plug-in's de Netscape

Los Plug-ins son programas que extienden las capacidades del visualizador Navigator de manera específica, por ejemplo la capacidad de mostrar vídeo, audio, archivos de un determinado formato (como: PDF, presentaciones de ASAP, archivo VRML, etc.).

No existe actualmente un conjunto estándar de plug-in para cada tipo de archivos, lo que existe son diversas aplicaciones, realizadas por diversos fabricantes de software y especialmente diseñados para Netscape Navigator, NO todos estos programas son de libre distribución, como es común en software para el Web, sin embargo se puede asegurar que todas las aplicaciones serán compatibles dentro de pocos meses.

Para que un archivo multimedia sea visible por el cliente debe cumplir dos condiciones: la primera es que posea el visualizador Navigator en su versión 3.0 o superior y la segunda es que tenga el plug-in correspondiente que trata el archivo deseado. Por tanto es recomendable un uso moderado de los plug-in, utilizando archivos que sean de formatos comunes y un uso complementario en una página, ya que muchos usuarios no podrán visualizarlo.

### <EMBED> insertar archivos

Esta etiqueta permite incluir cualquier tipo de archivo en la posición que se dese del documento HTML. Su funcionalidad es similar a la **IMG** para incluir imágenes. En este caso para que pueda mostrarse el archivo, se deberá tener el plug-in adecuado que lo interprete.

El formato es el siguiente:

```
<EMBED SRC="URL del archivo" WIDTH=n ó n% HEIGHT=n ó n%>
```

En el atributo **SRC** se indica el archivo que se desea mostrar, y los atributos **WIDTH** y **HEIGHT**, indican el tamaño que tendrá en objeto insertado, siendo estos opcionales.

## Scripts

Un Script es un programa escrito en un lenguaje de programación, el cual se puede poner en una página HTML. Para incluir este tipo de programas en una página, se debe utilizar la etiqueta **<SCRIPT>** **</SCRIPT>**. Esta etiqueta tiene el atributo **language** para indicar el lenguaje de programación utilizado por el Script. El código del programa debe "ocultarse" con las etiquetas de comentario de HTML: **<!-- -->** con el objeto de que no sean mostradas por los visualizadores que no reconocen la etiqueta **<SCRIPT>**. Los scripts más usados suelen estar escritos en **JAVA**. Este lenguaje permite realizar operaciones de animación y multimedia en una página Web. Así mismo se puede incluir con esta etiqueta programas escritos en C++, JavaScript, PERL o Visual Basic entre otros.

La inclusión de un Script en una página Web se hace de la siguiente manera (para el caso de un programa escrito en JavaScript):

```
<SCRIPT language = "JavaScript">  
<!--instrucciones del programa...-->  
</SCRIPT>
```



## Capítulo V

### **Caso de Aplicación:** Sitio Web del Departamento de Productos Interactivos para la Docencia. DGSCA

---

La Dirección General de Servicios de Cómputo Académico (DGSCA), es el centro de operaciones en materia de cómputo y logística informática más importante de nuestra Universidad. Entre sus funciones se cuentan las de desarrollar e instalar sistemas de cómputo en diversas dependencias universitarias, supervisar su correcto funcionamiento, ofrecer servicios de asesoría y capacitación a los miembros de la comunidad universitaria y al público en general. La DGSCA participa activamente en el *Comité Asesor de Cómputo*, proponiendo y discutiendo políticas institucionales de computación, informática y electrónica digital, además coadyuva en la supervisión del ejercicio presupuestal, así como en la compra y mantenimiento de equipo de cómputo.

Por otra parte, para impulsar la formación de recursos humanos, atiende un número considerable de estudiantes y profesores cada año, ofrece cursos de capacitación técnica, planes de becas, programas de actualización, diplomados, educación a distancia, etc. Actualmente se encuentra modernizada la red telefónica de la UNAM y mediante el proyecto denominado *Red Integral de Telecomunicaciones*, están funcionando seis estaciones de satélite, dos de ellas con servicio telefónico, y cinco estaciones de microondas, dando así soporte a Red-UNAM.

En materia de servicios de cómputo, maneja grandes equipos que proporcionan apoyo académico, participa en la organización de la *Comisión de Supercómputo (CRA Y)*, en la que se aportan ideas y soluciones encaminadas al desarrollo de la investigación científica.

Para lograr todo esto, la DGSCA cuenta con las siguientes Direcciones ubicadas en Ciudad Universitaria:

- Dirección de Telecomunicaciones Digitales.
- Dirección de Cómputo para la Investigación.
- Dirección de Cómputo para la Docencia.
- Dirección de Cómputo para la Administración Académica.

La Dirección de Cómputo para la Docencia tiene a su cargo la *difusión de la cultura informática y enseñanza en cómputo*, proporcionando una valiosa herramienta que complementa y enriquece la formación profesional a través de varios programas y proyectos, entre los que destacan:

- **Programa de Capacitación y Actualización en Cómputo:**  
Ofrece cursos de actualización y capacitación en cómputo orientado a capacitar universitarios y público en general en tecnología informática.
- **Diplomados:**  
En cómputo y telecomunicaciones, son con el objetivo de preparar y capacitar a personal en algunos de los más novedosos campos de estas disciplinas.
- **Líneas de Especialización:**  
Estas especializaciones brindan al estudiante sobresaliente, una secuencia lógica de cursos, que permiten profundizar en alguno de los campos en cómputo y telecomunicaciones modernas.
- **Educación a Distancia:**  
Este programa reúne alternativas tecnológicas de comunicación para la docencia, cuenta con recursos físicos, lógicos y humanos para llevar a la práctica talleres, seminarios, cursos y diversas actividades académicas a distancia (Videoconferencia y Audioconferencia), a nivel nacional e internacional.

Para ello cuenta con el apoyo de las de las direcciones que conforman la DGSCA, respaldadas por un cuerpo de académicos, profesionistas, estudiantes y técnicos en varios campos de investigación y aplicación específicos e diversas áreas de cómputo, así como de la infraestructura general que ofrece la dependencia universitaria.

La Dirección de Cómputo para la Docencia cuenta con diversos departamentos, para dar soporte a la gran diversidad de opciones que ofrece:

- Departamento de Control Escolar.
- Departamento de Relaciones y Mercadotecnia.
- Departamento de Desarrollo Curricular.
- Departamento de Infraestructura y Prospección Tecnológica.
- Departamento de Productos Interactivos para Docencia.
- Coordinación de Educación a Distancia.

De reciente creación (en verano de 1996), el **Departamento de Productos Interactivos para la Docencia**, a cargo del *Fis. Manuel López Michelane*, pretende dar soporte a los departamentos antes mencionados (dentro de la Dirección de Cómputo para la Docencia); llevando a cabo un *análisis, diseño, desarrollo, e implementación de sistemas (o "productos") interactivos de información*, en arquitecturas abiertas y de tecnología avanzada, en apoyo directo a las diversas actividades de esta dirección.

Al mismo tiempo se pretende difundir por Internet y específicamente a través de WWW, una serie de *prototipos de cursos de computación*<sup>1</sup>, empleando para ello las características hipertexto del lenguaje HTML y la infraestructura con la que cuenta el departamento.

El Departamento de Productos Interactivos para la Docencia cuenta con la siguiente infraestructura para el desarrollo de aplicaciones en el Web: una estación de trabajo Silicon Graphics "Indy" R4600pc a 133Mhz, con un disco duro interno de 512 MB, 32 MB RAM, este equipo posee instalado entre una gran variedad de aplicaciones para Internet, un *software-servidor* (NetSite Communication Web Server) de NCSA, bajo un sistema operativo UNIX (IRIX IV) también de Silicon Graphics; la estación de trabajo tiene conexión (o salida) a Internet a través de RedUNAM, con el siguiente URL asignado: <http://morsa.dgscsa.unam.mx>. Tanto hardware como software, se encuentran a disposición del departamento y actualmente operando como Servidor Web de este.

Existe también un equipo de trabajo (humanware), conformado por: ingenieros, analistas, diseñadores gráficos, personal de soporte, becarios, etc., al cual me encuentro integrado con el nivel de *Becario-Asistente* en apoyo directo al proyecto de creación del Sitio Web del departamento, **caso de aplicación demostrativo de este trabajo de Tesis:**

***El objetivo que pretende alcanzar el departamento es el siguiente:***

*Tener un espacio electrónico o Sitio Web del Departamento de Productos Interactivos para la Docencia, el cual sirva para evaluar la posibilidad de proporcionar a mediano plazo material de apoyo didáctico (apuntes, manuales, tutoriales, etc.), e información en línea que sirva primero, como material de consulta y posteriormente como curso interactivo, empleando la gran cobertura de Internet y difusión a través de WWW.*

Para ello se deben lograr las siguientes *objetivos específicos*:

- *Difundir a la comunidad universitaria y público en general de habla hispana, información de cómputo con carácter de libre distribución, a través de WWW.*
- *Proporcionar pequeños prototipos de cursos de computación, tales como: tutoriales, manuales y apuntes en línea (online), en idioma español mediante el empleo del hipertexto.*
- *Conocer que tipo de usuarios finales (o clientes), se interesan por la información que puedan obtener del servidor Web, así como su procedencia.*
- *Saber la opinión de los usuarios respecto a los prototipos expuestos, con la finalidad de mejorar su alcance y contenido para su posterior liberación y actualización, una vez efectuadas las revisiones correspondientes (por parte de los diversos Colegios de la DGSCA).*

<sup>1</sup> Estos cursos de computación son *prototipos aún No liberados*, el contenido de cada uno de estos, se encuentra en aprobación por parte de los diversos colegios que integran los órganos revisores para cada disciplina específica que se imparte en la DGSCA, por ejemplo existe: El Colegio de Bases de Datos, Colegio de Internet, Colegio de Introducción a la Computación, etc.

## **Análisis del público y fuentes de información**

El conjunto de páginas Web está enfocado a la difusión de la enseñanza de cómputo, dirigido principalmente a jóvenes universitarios y *comunidad Internet de habla hispana*, que deseen expandir sus conocimientos en Informática. La información que se presenta tiene la finalidad de orientar al usuario para que posea una idea clara del tema expuesto.

Una de las tareas más caras y lentas en la organización de un sistema para la difusión y promoción de información como es el Web, es la recopilación del conjunto de los datos que conforman el contenido de esta, no sólo por la búsqueda constante debido a que los temas a tratar son sobre cómputo, sino por vigilar el buen contenido y redacción de la información.

Sin embargo, la necesidad de contar con un espacio en WWW por parte del Departamento de Productos Interactivos, orilló a la puesta inmediata en el servidor Web de los prototipos de cursos a difundir, de esta manera el contenido, los guiones, así como la redacción se encuentran aún en revisión por parte de los diversos Colegios de la dependencia (invierno de 1996).

Para obtener los datos sobre los temas de interés, se consultaron a: académicos, técnicos especialistas y becarios entre otros, apoyados fundamentalmente en material de consulta pública editado por la DGSCA, y de una vanada bibliografía y apuntes recopilados en varios departamentos.

Cabe mencionar que la totalidad de la información, productos que se generen a partir de este y material de consulta contenido en el servidor del departamento, serán protegidos por los derechos de autor correspondientes a la DGSCA, en su carácter de dependencia universitaria.

## **Diseño e Implementación**

Para el diseño de páginas en WWW, no existe una metodología específica a seguir, particularmente porque el proceso de diseño, al igual que los previos de planeación y análisis, se ven influenciados directamente por el tipo de información específica que se desea difundir, así como de la complejidad de la aplicación y la forma como se desea estructurar e integrar dicha información<sup>2</sup>. Este proceso puede continuar incluso después de que la aplicación haya sido liberada; sin embargo, se deben considerar dos planteamientos:

---

<sup>2</sup> Para mayor detalle consulte el apéndice A "Metodología del Desarrollo", en el se exponen las etapas de Planeación, Análisis, Diseño, Implementación, Promoción e Innovación para WWW, extraídas del Capítulo II de la tesis "Uso de HTML y CGI para el desarrollo de aplicaciones en Internet" de César Cruz Cortés.

- **Diseño top-down.** Se emplea cuando se tiene una idea clara *a priori* del sistema, ya que se comienza con una página inicial o frontal llamada *página base* o *home page*, a partir de la cual se desprenderán la totalidad de páginas que conforman el sistema.
- **Diseño bottom-up.** Es útil cuando No se conoce la apariencia final que deberá tener el sistema de páginas, pero se sabe claramente cómo deben de verse y las funciones específicas que debe cumplir cada una de ellas; de esta manera, se procede a unir las páginas específicas en estratos superiores que las contengan, hasta llegar a conformar la *página base* del sistema.

La unión o integración de páginas para los dos casos, se pueden implementar mediante *árboles de navegación*, estos proporcionan una idea general de las conexiones entre las diferentes áreas que componen el sistema, además de ayudar a organizar el contenido, formato y mensajes durante la visualización de las páginas. Básicamente las estructuras para desarrollar un árbol de navegación en una aplicación Web son:

- **Lineal:** Es donde se navega de modo secuencial, de una página de información a otra.
- **Jerárquica:** Es donde el usuario navega a través de las diversas ramificaciones de una estructura árbol, que se forma dada la lógica natural del contenido.
- **No lineal:** El usuario navega libremente a través del contenido del sistema, sin limitaciones de visualización.

Gran parte del sistema desarrollado es No lineal, ya que por la naturaleza del tipo de publicación Web, los clientes deben tener la libertad de saltar de un índice, de un gráfico o de un comentario a otro, a través de la flexibilidad que permite el hipertexto (implementado en HTML). El tipo de estructura empleado para la difusión de la aplicación combina varios modelos de consulta sobre una estructura jerárquica, adecuada para el tipo de información mostrada (cursos prototipo), con opciones de consulta no lineal y con segmentos lineales que conforman la explicación de los temas de cómputo. Sin embargo, es posible cambiar a un modo no lineal en el que el usuario pueda consultar cualquier tema sin importar en que parte del sistema de páginas se encuentre.

Los temas incluidos hasta ahora (invierno de 1996), fueron seleccionados al azar, ya que como prototipos, pretendían evaluar la posibilidad de apoyar a las diversas actividades docentes de la dependencia. La intención fue realizar un sistema funcional, que sirva como base para el desarrollo de aplicaciones más detalladas, al cual se integren nuevos conocimientos para difundirlos libremente (sin costo alguno).

Para el caso de la aplicación desarrollada se efectuó primero un diseño bottom-up, con la asignación respectiva para crear las diversas páginas que integrarían temporalmente el conjunto de los cursos prototipo, los cuales pueden ser localizados en el siguiente URL:  
<http://morsa.dgsca.unam.mx>

Un índice de los productos disponibles al final de diciembre de 1996 se puede ver en la siguiente figura:

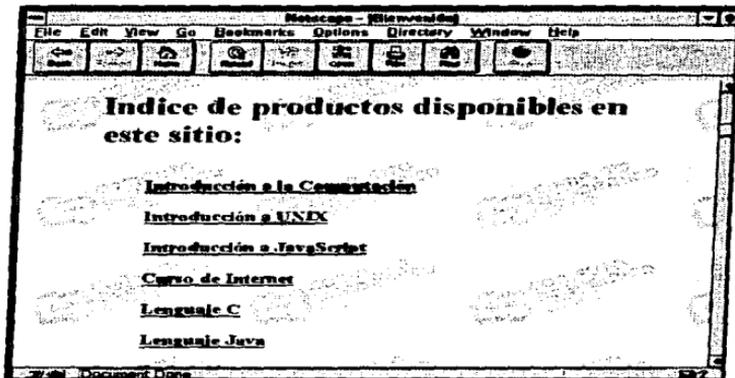


Figura 5a Índice de Prototipos disponibles a finales de 1996.

El código HTML es el siguiente, contenido en el archivo indice.html, ubicado en el directorio `usr/local/etc/httpd/htdocs/indice.html`, (directorio asignado por omisión durante la instalación del software de servidor Web):

```
<html>
<head><title>Bienvenido</title></head>
<body text=#000000 vlink=#23238e background="grafic/dgsca20.gif">
<center>
<table width=80% border=0><tr><td>
<font color=236b8e><h1>Indice de productos disponibles en este sitio:</h1></font><p><br>
</center>
<ul><font size=+1 color=#0000ff><br>
<p>Introducción a la Computación
<p>Curso de Internet
<p>Introducción a JavaScript
<p>Lenguaje Java
<p>Evaluación (prototipo)</font>
</font></ul><p><br>
```

```

<center>
<table border=0><tr><td>
<font size=+1>S</font>i tiene alguna duda o comentario acerca de esta p&aacute;gina
compartala con nosotros, envienos un <b><a href="mailto:ulises@morsa.dgsca.unam.mx">
correo electr&ocute;nico</a>.</b></td>
<td><a href="mailto:ulises@morsa.dgsca.unam.mx"></a>
</td></tr></table><p><br><p>
Ultima actualizaci&ocute;n: 1 de diciembre de 1996<p><br><p>
<i><font size=2 color=#23238e>Copyright &copy; 1996 by</font><br>
<b><font color=#236b8e> Productos Interactivos para la Docencia<br>
Direcci&ocute;n General de Servicios de C&ocute;mputo Acad&ocute;mico<br>
Universidad Nacional Aut&ocute;noma de M&ocute;xico.</font></i></b>
</td></tr></table>
</center>
</body>
</html>

```

Una vez efectuada la integraci&ocute;n de los prototipos, se tom&ocute; la decisi&ocute;n de llevar a cabo un dise&ocute;no top-down con la incorporaci&ocute;n de una p&ag;ina base o "home page":

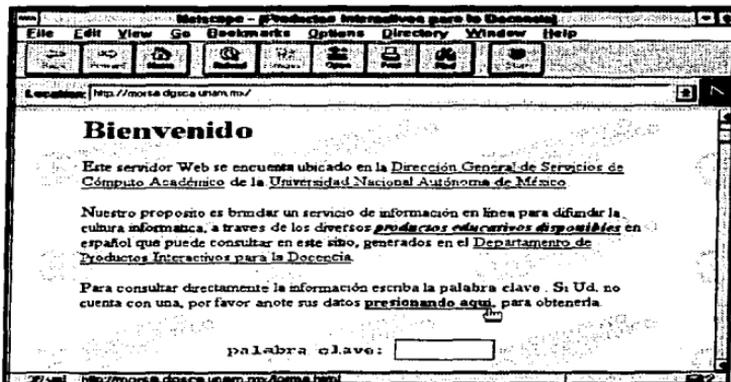


Figura 5b P&ag;ina base del Departamento de Productos Interactivos para la Docencia.

<http://morsa.dgsca.unam.mx>

## Creación de Páginas en WWW

El código HTML de esta página es el siguiente:

```
<html>
<head><title>Productos Interactivos para la Docencia</title></head>
<!-- bgcolor=#000000 alink=# link=# -->
<body text=#000000 vlink=#23238e background="grafic/dgsgca20.gif">
<FORM METHOD="POST" ACTION="http://morsa.dgsgca.unam.mx/cgi-bin/informacion.pl">
<center>
<table width=85% border=0><tr><td>
<b><h1><font color=#236b8e>Bienvenido</font></h1></b><p>
<font size="+1">E</font>ste servidor Web se encuentra ubicado en la <a
href="http://www.museovirtual.unam.mx/dgsgca.html">Dirección</a> de la <a
href="http://www.unam.mx">Universidad
Nacional Autónoma de México</a>. <p>
<font size="+1">Nuestro propósito es brindar un servicio de información; en
lenguaje sencillo para difundir la cultura informática, a través de los diversos
<i><b><a
href="productos.html">productos educativos disponibles</a>
</b></i> en español que puede consultar en este sitio,
generados en el <a href="http://depidoc.html">Departamento de Productos Interactivos para la
Docencia</a>. <p>
<font size="+1">Puede consultar directamente la información; escriba la <b><font
color=#236b8e size=3>palabra clave</font></b>.
<font size="+1">Si Ud. no cuenta con una, por favor anote sus datos <b><a
href="forma.html">presionando aquí</a>, para obtenerla.
<br><p><br>
<center>
<pre>
<p><b><font color=#236b8e><blink><font size="+1>palabra clave:</b>
<font><blink></font><INPUT TYPE="password" name="password" SIZE=10<br>
<p>
<b><i><font size=3>Disculpe las molestias que esto le ocasiona.</font></i><br>
<INPUT TYPE="submit" VALUE=" Entrar "> <INPUT TYPE="reset" VALUE=" Limpiar
"><br>
<p></a></pre><br>
<font size="+1>Si tiene alguna duda o comentario, envíenos un
<a href="mailto:ulises@morsa.dgsgca.unam.mx">correo electrónico</a>. <p><br><p>
</td></tr></table>
<p><br><p>
Ultima actualización: 1 de diciembre de 1996. <p><br><p>
<i><font size=2 color=#23238e>Copyright &copy; 1996 by</font></i><br>
<b><font color=#236b8e> Productos Interactivos para la Docencia</b>
```

Director General de Servicios de Computo Académico  
 Universidad Nacional Autónoma de México

</center>  
 </body>  
 </html>

*No es posible proporcionar en forma impresa (a través de este trabajo de Tesis), la totalidad del conjunto de páginas que integran el sitio Web del Departamento de Productos Interactivos para la Docencia, ya que como se puede observar, para el caso ilustrativo de dos páginas, sin demasiadas etiquetas del lenguaje HTML se tiene un código extenso.*

*Debido a la naturaleza de constante actualización e innovación que presentan este tipo de aplicaciones, mostrar el creciente y variado número de páginas ocuparía demasiado espacio impreso; sin embargo, gracias a las capacidades y flexibilidad de la gran mayoría de los visualizadores modernos se puede obtener una copia fiel del documento, mediante los diversos menús de edición proporcionados para este fin (por ejemplo: "save as", disponible en muchas plataformas gráficas).*

*Cabe destacar que la información contenida en el servidor Web del departamento, hasta hoy es de dominio público (primavera de 1997), por lo tanto cualquier persona es libre de consultar la información y disponer de ella exclusivamente en forma personal para enriquecer su cultura informática. (No esta permitida la reproducción total o parcial de la información contenida en el servidor: <http://morsa.dgscs.unam.mx> para fines lucrativos, tampoco se permite usar este material en una institución de enseñanza o usar extractos para una publicación de cualquier tipo, para ello debe solicitar permiso previo y por escrito de sus autores).*

Finalmente durante algunas revisiones hechas al sistema, se pudieron observar ciertos factores a considerar como son:

- **Comportamiento del usuario.**

El Web es un sistema dinámico basado en selecciones del usuario, por lo que no hay forma en que un desarrollador pueda controlar la manera en que se accederá y empleará la información de la aplicación. A pesar de que el comportamiento del usuario no puede ser controlado, éste puede ser canalizado para intentar "guiarlo" a través del sistema al gusto del programador (mediante ligas hipertexto únicas a páginas definidas); *sin embargo la última palabra la tiene el usuario final.*

- **Visualizadores.**

Debido a la naturaleza cliente-servidor del Web, es imposible conocer el visualizador-cliente que el usuario estará empleando para ver la aplicación. Como se menciona en el capítulo 2, existen numerosos visualizadores o "navegadores"; durante el desarrollo de la aplicación se debe considerar qué información será esencial, de manera que nada se pierda, cuando se empleen visualizadores con capacidades inferiores a las esperadas (es decir, que no muestren gráficos, o bien, que no soporten algunas de las extensiones HTML).

- **Ligas dentro y fuera del sistema.**

En cualquier aplicación Web podemos encontrar ligas que hacen referencia a recursos de WWW e Internet, totalmente fuera del sistema desarrollado; estas ligas se encuentran aisladas, fuera de control por parte del desarrollador de la aplicación, así que al incluirlas en ella, uno podría encontrarse con que la liga hace referencia a páginas con información no actualizada, o que ha cambiado debido a las disposiciones de sus autores.

# Conclusiones

---

- El presente trabajo fue generado en base a la necesidad de tener un Sitio Web en el Departamento de Productos Interactivos para la Docencia, el cual, gracias a las propuestas e implementaciones efectuadas, difunde actualmente una variada cultura informática (constantemente actualizada). A través del material expuesto a la comunidad universitaria y público en general de habla hispana, se han logrado obtener óptimos resultados como: el apoyo didáctico que presta a cursos que se imparten en la Dirección General de Servicios de Computo Académico, material de apoyo contenido en la dirección: <http://morsa.dgscs.unam.mx>. Así también, la publicación del presente trabajo constituye una guía que sirve para el desarrollo de aplicaciones en el Web; cabe señalar que los capítulos III y IV proporcionan un manual en español de lenguaje HTML con variados ejemplos.
- La aplicación final, se apega plenamente a las políticas que maneja la Universidad Nacional Autónoma de México, a través de la DGSCA, en su interés constante por difundir el conocimiento humano; tomando en cuenta que la ingeniería es una disciplina que se dedica a la solución de la problemática que afecta al individuo y su sociedad, apoyada siempre en el estudio y conocimiento de las leyes que rigen la naturaleza; a este conocimiento teórico se agregó una visión práctica, en búsqueda de un mayor provecho de los recursos disponibles, de esta manera, se puede afirmar con certeza que *en base a los recursos plenamente identificados, requerimientos definidos y al conjunto de soluciones implementadas, se presenta una solución real a las necesidades planteadas al inicio de esta tesis.*
- El Web es un conjunto de conocimientos que cambia en forma constante su contenido, debido a la creciente necesidad por *naturaleza humana*, de estar actualizado en cuanto a información se refiere, ayuda a la gente a conocer temas tan variados que van desde tópicos como medicina, teatro, pasando por política o costumbres y religión, hasta exploración espacial y noticias. WWW es un medio para la difusión del conocimiento humano, el cual posee un formato dinámico que permite una comunicación masiva y a la vez personal. Gracias a la facilidad de su uso e implementación, comparado con otros sistemas de recuperación de información en Internet como: Archie, Gopher o WAIS, este medio de comunicación enlaza a la gran mayoría de recursos, servicios y aplicaciones que existen en la "red de redes". Su éxito radica en la integración que presenta de variados formatos de información como: imágenes, texto, e inclusive audio y video.

## Conclusiones

- Gracias a la capacidad hipertexto e hipermedia del lenguaje HTML y el empleo del conjunto de protocolos TCP/IP, WWW ha resultado ser una fórmula exitosa, dejando de ser uno más de los servicios disponibles en la "red de redes", para convertirse en el impulsor del explosivo crecimiento y difusión de Internet en todo el mundo.
- HTTP es la base operacional de WWW, este protocolo de solicitud/respuesta ubicado en la capa de aplicación, lleva instrucciones y código de las peticiones de los visualizadores hasta los servidores Web, y las respuestas de estos de regreso a los clientes o visualizadores. Al continuar el explosivo crecimiento de WWW y agregar nuevas características tanto a clientes como a servidores, un *protocolo de transferencia estandarizado* como HTTP, es esencial para conservar las crecientes funcionalidades del Web y su interoperabilidad.
- Es necesario identificar las necesidades de mantenimiento de la información en cualquier tipo de aplicación para WWW, siendo este un sistema de información de rápida y fácil actualización, por lo que resulta indispensable una actitud dinámica, que permita a los usuarios tener la certeza de que los datos proporcionados por la aplicación, son lo suficientemente recientes y verídicos como para confiar en ellos; para lograrlo, se proponen ciertos consejos y una pequeña guía de estilo ubicada en el Apéndice A, además de esto, propongo las siguientes consideraciones, para ser tomadas en cuenta durante el proceso de creación, puesta en marcha y buen funcionamiento de páginas en WWW:

De ser posible, si se desea tener más servicios de información en un departamento específico, emplear un equipo de cómputo para cada grupo de servicios, por ejemplo: WWW y Gopher; de no ser posible, hay que distribuir los puertos de comunicación de los servicios con mayor acceso, para evitar saturación

Se debe recordar que el ancho de banda es una limitante para el desarrollo de aplicaciones que requieren transferir grandes cantidades de información a muchos puntos diferentes, hay que considerar opciones que brindan lenguajes de programación desarrollados para Internet, como JAVA y JavaScript, los cuales trasladan ciertas actividades de los servidores a los clientes-visualizadores para ser procesadas, evitando el congestionamiento en el servidor.

Las páginas en el servidor Web deben estar organizadas por temas, no excediendo las capacidades de despliegue en cuanto a un número máximo de menús, los cuales deberán ser proporcionales al contenido en donde se alojan, así como procurar tener una referencia breve de la información que contienen.

Para cada tema en específico o servicio proporcionado, se debe tener una cuenta de correo electrónico, a la cual se puedan enviar comentarios y sugerencias para mejorar el contenido y alcance de la aplicación mostrada, hay que recordar que hablamos de un público de miles de hispano hablantes en WWW.

No hay que emplear demasiadas imágenes, ya que esto genera un documento grande, lo cual implica una transferencia de información lenta, con la consecuente distracción o abandono del sitio Web, o bien utilizar imágenes pequeñas en formatos que permitan ahorrar espacio.

WWW se encuentra hasta cierto punto saturado de basura, información que es redundante, inconsistente, la cual no proporciona conocimiento alguno, para evitar esto, hay que depurar el contenido a fondo y evaluar la permanencia de esta en el Web.

Es primordial efectuar respaldos continuos de la información en el servidor Web, sobre todo si no se cuenta con un dispositivo de seguridad (como un firewall), recordemos que lo más importante de un servicio es la información.

Por último, hay que mantener bitácoras de todos los servicios proporcionados, y revisarlas periódicamente para conocer la afluencia, intereses y preferencias de los usuarios, y en base a esto aumentar la capacidad de disco, procesador o memoria del servidor Web.

- Actualmente como se puede observar, la fusión entre las tecnologías de telecomunicaciones y computación, específicamente refiriéndose a Internet y World Wide Web, son tecnologías que están revolucionando nuestra forma de vivir, expandiendo el conocimiento humano a todos los rincones de la Tierra, sin censuras o restricciones de algún tipo, se trata de una tecnología *"fuera de control"*, en donde no se conoce con exactitud a donde conducirán al hombre. *"La red"*, que comenzó como un proyecto de comunicación ante la posibilidad de un ataque nuclear, ha derivado en una gigantesca *"red de redes"*, que comunica de manera amplia y eficiente a un gran y creciente número de personas alrededor del mundo; generado una nueva cultura, la *"cybercultura"*, con su modo de hablar, pensar, sentir e imaginar; un mundo nuevo en el que aún falta mucho por explorar y que tiene un gran potencial. Probablemente este sea solo el comienzo de una nueva era en la forma de comunicarnos: *"unos"*, con *"otros"*, de la cual somos partícipes dentro de un episodio más en la historia de la humanidad.

**Ulises Cruz Cortés.**

## ***Conclusiones***

# Apéndice A

## Metodología de desarrollo

---

Los procesos comprendidos para el desarrollo de aplicaciones en WWW se encuentran íntimamente relacionados, esto son: Planeación, Análisis, Diseño, Implementación, Promoción e Innovación.

### Planeación:

La planeación es un aspecto importante en el desarrollo de aplicaciones, pues es aquí donde son tomadas decisiones que afectan al diseño, implementación y posterior promoción de la aplicación.

La planeación por lo regular se presenta dentro de un contexto más general y amplio que el de la simple identificación de componentes técnicos de un conjunto de páginas HTML; frecuentemente —en especial para grandes organizaciones— la comunicación a través de WWW es parte de un esfuerzo estratégico para alcanzar a los consumidores.

El primer paso en la planeación de un sistema Web es **definir** la manera en que éste pudiera jugar un papel en las necesidades de comunicación de la organización. Este proceso de definición puede comenzar con la evaluación de los medios de comunicación que ya se estuvieran empleando (como: radio, televisión, prensa, etc.); la aplicación no necesita reemplazar todos los métodos existentes, sino complementarlos.

Después de la definición sigue la **integración**, donde la aplicación Web deberá relacionarse con la infraestructura de comunicación ya existente en la organización (si es que existe), por ejemplo: algún servidor ftp ó gopher que también sean empleados como medio de comunicación del organismo. Es decir, se tratar de elaborar un plan para "ligar" elementos de la aplicación a flujos de información ó comunicación ya establecidos.

Finalmente, la **diferenciación** es donde la aplicación Web debe mostrar ventajas claras sobre algunos de los medios tradicionales de comunicación de la organización, o de otra forma será necesario detener el desarrollo del sistema. La tecnología de punta implicada en WWW no significa que las aplicaciones Web sean la solución a todas las necesidades de cualquier organismo.

Para el caso particular del Web se deben considerar los siguientes puntos:

- **Información de la Audiencia.** Para que exista una comunicación precisa, es necesario que un desarrollador sepa lo que se quiere comunicar y a quién. La información que se tenga del "destino final" ó auditorio, es de vital importancia para el establecimiento de una comunicación exitosa con el usuario, asegurando una mejor transmisión de los mensajes o información que se desea dar a conocer.
- **Establecimiento del objetivo.** Establecer un objetivo general sirve como guía a lo largo de todo el desarrollo de la aplicación; sin un objetivo, los analistas no cuentan con ninguna base para evaluar el sistema y decidir si éste trabaja eficientemente o no.
- **Dominio de la información.** Es necesario tener dominio de la información que habrá de emplearse para el desarrollo del Web; deberá incluirse la que se proporcionará al usuario, así como la que los desarrolladores deberán conocer para realizar un buen trabajo. Además, deberán hacerse consideraciones acerca de las formas en que ésta será obtenida, así como de planes para la actualización de la información que necesite ser actualizada continuamente.

## **Análisis:**

Este proceso contempla desde la validación técnica de una implementación HTML, hasta el estudio del contenido y diseño de una aplicación ya existente o bien, que apenas se encuentre durante la etapa de planeación.

La información sobre la que se centra éste proceso, no es únicamente la que se ofrece al usuario en la aplicación, sino también se refiere a la información obtenida por el equipo de desarrollo.

Para el caso del Web se deben considerar los siguientes puntos:

- **Desempeño.** En donde se deberá contemplar los siguientes puntos: *Tiempo de retardo*, se refiere al tiempo de espera para visualizar una página de la aplicación. *Legibilidad*, se debe procurar que los recursos empleados para "adornar" la aplicación (como: gráficos, texturas, etc.), no entorpezcan la legibilidad del documento, y la *disponibilidad de uso*, es decir, que la información generada esté disponible para una variada gama de clientes-visualizadores.
- **Estética.** Este aspecto es difícil de evaluar, sin embargo, pueden emplearse ciertos lineamientos como un diseño balanceado y coherente que ayuden al usuario a enfocar su atención en el contenido de la aplicación.
- **Uso.** Durante el análisis se deberá verificar la facilidad de uso de la aplicación, así como la utilidad de la información que ésta proporciona al visitante.

## **Diseño:**

Durante esta etapa se deben crear lineamientos para la implementación de la aplicación, que tengan elementos suficientes para que el usuario obtenga información al nivel deseado. Para este tipo de diseño, NO existe una metodología específica a seguir, particularmente porque este proceso, al igual que los anteriores, pueden continuar incluso después de que la aplicación haya sido liberada. Sin embargo, se deben considerar dos tipos de diseño:

- **Diseño top-down.** Se emplea cuando se tiene una idea clara *a priori* del sistema, ya que se comienza con una página inicial ó frontal llamada *página base ó home page*, a partir de la cual se desprenderán la totalidad de páginas que conforman el sistema.
- **Diseño bottom-up.** Es útil cuando No se conoce la apariencia final que deberá tener el sistema de páginas, pero se sabe claramente cómo deben de verse y las funciones específicas que debe cumplir cada una de ellas; de esta manera, se procede a unir las páginas específicas en estratos superiores que las contengan, hasta llegar a conformar la *página base* del sistema.

Existen ciertas técnicas que ayudan a considerar aspectos que son importantes para el manejo y buen despliegue de la información, estas son:

- **Agrupamiento de la información en bloques de un tamaño adecuado.** Se trata de agrupar moderadamente la información desplegada al usuario, la cual no deberá mostrar mucha información en una sola página, en caso de ser necesario se debe hacerse uso de la *capacidad hipertexto de WWW*.
- **Establecer una apariencia uniforme en las páginas que componen la aplicación.** Es importante que las páginas que integran un sistema, posean un formato o patrón uniforme que haga sentir al usuario final, la pertenencia de cada una de ellas a un sistema global.
- **Creación de páginas índices.** En aplicaciones grandes, compuestas de una gran cantidad de documentos, se requiere de páginas índice es decir, páginas con ligas hipertexto que guíen al usuario a lugares o sitios específicos con la información disponible en todo momento.

## **Implementación:**

En esta etapa se traducen todas las especificaciones arrojadas por los procesos anteriores en código HTML. En caso de que el sistema de páginas Web interactúe con alguna otra aplicación externa, por ejemplo mediante CGI, es en esta etapa cuando también se genera el código correspondiente el cual a su vez obedece al conjunto de lineamientos establecidos por las etapas anteriores del proceso de desarrollo.

Para llevar a cabo una implementación adecuada, es recomendable considerar los siguientes puntos:

- **Trabajo en conjunto.** Se debe dedicar tiempo al trabajo con otras personas: planeadores, diseñadores, analistas, promotores, usuarios, etc.; la finalidad es simplemente poder establecer vías de comunicación entre ellos, la cuales permitan un desarrollo fácil y apegado a los objetivos iniciales de la aplicación.
- **Fase de pruebas.** Una vez escrito el código de la aplicación, se obtiene un *prototipo*, el cual debe ser sometido a pruebas para verificar si puede ser "liberado", o en su defecto canalizado al proceso de desarrollo correspondiente para ser corregido.
- **Mantenimiento continuo.** Se debe dar un mantenimiento constante el cual puede nunca terminar, pues el crecimiento y cambios son características inherentes de las buenas aplicaciones Web.

## **Promoción:**

Una vez finalizada la aplicación es necesario darla a conocer, empleando para ello la misma red global o algún otro medio de comunicación disponible, utilizando técnicas que permitan alcanzar objetivos como: informar al público en general de la existencia de la aplicación así como su oferta, atraer el interés de la audiencia planeada y dar a conocer la manera en que el sistema satisface sus necesidades de información. Existen vías de comunicación útiles que se pueden emplear para publicitar un sistema tales como: medios de comunicación existentes, servicios en línea, promoción mediante administradores de otros sitios con fines comunes o patrocinio por parte de alguna compañía o institución y anuncios clasificados.

## **Innovación:**

La innovación es un proceso creativo y dinámico que no puede ser comprendido en una serie de pasos a seguir; implica un constante monitoreo y conocimiento de las necesidades de los usuarios conforme éstas vayan surgiendo, así como el desarrollo de estructuras que nos permitan conocer dichas necesidades, para ello se debe tener un monitoreo de las necesidades de información. La idea central de la innovación es dejar correr la aplicación Web de acuerdo a lo previsto, dentro del sistema deberán existir medios por los cuales, los usuarios puedan comunicarse con los administradores y así identificar las nuevas necesidades de información de los usuarios, posibilitando la modificación de la aplicación de acuerdo a las nuevas condiciones y requerimientos de uso que deseen satisfacer.

Por último es necesario vigilar la calidad de la información arrojada por la aplicación, la cual deberá ser correcta, accesible a cualquier usuario, útil y entendible, así como considerar los avances tecnológicos para su incorporación en la aplicación.

## **Guía de Estilo Consejos de Diseño**

Conforme WWW sigue creciendo y se diversifica, es conveniente crear documentos HTML de acuerdo a "ciertas reglas" (NO definidas ni estandarizadas), sobre todo por la amplia diversidad de programas clientes WWW alrededor del mundo, es importante escribir documentos que se vean bien en cualquier visualizador y no en alguno en especial.

A continuación presento una guía de estilo con ciertas consideraciones de diseño, esta contiene sugerencias personales y sugerencias útiles, tales como prácticas a seguir, errores comunes y cosas que se deben evitar. Siguiendo estos consejos, quizás el documento no se vea tan bien como sería posible en un cliente WWW en particular, pero impedirá que se vea mal en cualquier otro programa cliente.

Por ejemplo, es bueno destacar quien o quienes son los autores de las páginas presentadas, y dotar a cada documento que se incluye en el servidor con los correspondientes créditos, de modo que el usuario pueda darse una idea de quien es la persona responsable de la presentación, así como de la compañía o institución a la que representa.

Los siguientes errores deben ser evitados:

### **• Errores al inicio de párrafos**

Este es, probablemente, el error más común y la causa número uno de que los documentos den una mala apariencia. La etiqueta <P> es como una "roptura" de párrafos, más que un terminador de estos. De acuerdo a la especificación HTML, <P> es usado entre dos piezas de texto los cuales de otro modo seguirían juntos. Funcionalmente, <P> sirve como un marcador de fin de párrafo, sin embargo en ciertos contextos el uso de <P> debe ser evitado, como directamente antes y después de cualquier otra etiqueta que implique una "ruptura" de párrafo, por ejemplo <LI>, <DT> y <DD>.

### **• Errores comunes de URL**

**Errores a referencias a directorios.** Al solicitar un documento, es común utilizar el formato: *servicio://maquina.dominio:puerto/ruta/archivo* en vez de *servicio://maquina.dominio:puerto/ruta/archivo/*. Algunos servidores son capaces de detectar éste error y proveen el URL adecuado, pero es mejor hacer la petición URL correcta, ya que no todos los servidores son capaces de soportar direccionamiento transparente.

**No usar nombres completos de dominios.** El problema surge cuando el nombre de host del URL no se encuentra perfectamente identificado por redes locales. Es decir uno puede hacer referencia a su propio servidor usando únicamente WWW localmente, pero el nombre correcto para el caso de la unam sería: "www.unam.mx".

**Uso inadecuado de URL's relativos.** Es posible crear un URL relativo, el cual posee la ventaja de ser más corto, y hacer que una colección de documentos sea más portable (fácil de moverse de un directorio a otro y de un servidor a otro). Un URL relativo es aquel que no contiene todas las partes necesarias de un URL (esquema, host, vía, información), este último se denomina URL absoluto. El visualizador intentará asumir que las partes faltantes se encuentran en el documento que inició la liga.

**Olvidar comillas.** Un error común al hacer una referencia es el olvidar el uso de las comillas (") en una etiqueta.

```
<A HREF="informatica.aragon.unam.mx">
```

**Etiquetas perdidas de cierre.** Muchas etiquetas HTML contienen información entre ellas, por ejemplo: <Strong>texto resaltado</Strong>. En muchos casos, la falta de una etiqueta de cierre puede causar efectos imprevistos al mostrar el documento.

### **Se debe evitar:**

- **Mezclar etiquetas del encabezado con elementos del cuerpo.**  
HTML especifica que no deben mezclarse elementos del encabezado <HEAD> con elementos del cuerpo <BODY>. EL resultado se refleja en la legibilidad del documento.
- **Evitar uso de espacio innecesario.**  
En general, el uso de espacios en blanco alrededor de etiquetas debe ser evitado.
- **Usar encabezados.**  
HTML establece que un encabezado no debe estar más allá de un nivel del encabezado anterior. Por ejemplo <H3> no deberá seguir a <H1>
- **Evitar enlaces sin significado.**  
Al generar un documento, se deberá asegurar que los enlaces tengan significado. Es decir, que el texto a ligar siga el flujo natural del tema que trata, especialmente hay que evitar el síndrome de "presione aquí" o "click here". (mucho software inicio, y subsiste, como visualizador de hipertexto sin orientación gráfica).

## **Consejos de Diseño:**

- **Páginas pequeñas**

El tener páginas muy grandes o extensas provee de mayor información al usuario, pero incomoda la búsqueda y revisión de ésta. Las páginas muy grandes consumen tiempo de transmisión, despliegue y mucha paciencia por parte de los visitantes. Se debe tratar de conservar lo más pequeño posible las páginas, inclusive dividiendo en partes el documento extenso, esto permite hacer un mejor uso de los recursos de Internet (además de que los usuarios lo agradecerán).

- **Informar sobre el enlace**

Las hiperligas o enlaces de un documento son muy importantes, es responsabilidad del diseñador proveer de información necesaria al usuario para que conozca de antemano que clase de documento visitará. Es necesario ser explícito en cuanto a los enlaces; un enlace textual debe hacer una referencia clara al contenido del documento a visitar o a la acción que va a realizar; un enlace gráfico debe ser lo más claro posible, algunas veces, una imagen por si misma explica todo, en otras, puede prestarse a interpretaciones erróneas.

- **NO incluir páginas muertas**

En sistemas con una gran cantidad de documentos es común el incluir páginas muertas, es decir páginas que no son localizables por el servidor de archivos y son presentadas al usuario como un error de recuperación. Son, también aquellas que, aun estando presentes, carecen de información referente al tema y conducen al usuario a un archivo "Bajo Construcción"; evítele perder la atención del usuario.

- **Consistencia en el documento**

Cada una de las páginas debe compartir las mismas etiquetas, la razón radica en la importancia de la uniformidad de la documentación. Un usuario común deben encontrarse con etiquetas con las que fácilmente identifique y que puedan invocar, comprender y analizar en cualquier documento que visite. Es importante evitar la inclusión de páginas que no compartan los mismos elementos en el servidor de archivos.

- **Evitar enlaces a gráficos**

El uso de gráficos en páginas de información se ha vuelto muy común. Muchos servidores mantienen enlaces de documentos o iconos hacia imágenes de alta calidad o muy grandes, algunos otros incluyen gráficos en documentos tomando la decisión del usuario a la elección de traerías a su programa cliente. Es necesario

evitar el enlace a gráficas en la medida posible, ya que, si no viene al tema simplemente será un adorno "bonito". Las imágenes muy grandes y de alta calidad consumen mucho tiempo de transmisión, NO hay que fatigar al usuario en la espera de imágenes que bien pueden ser suplidas por documentos texto. Se debe tener en cuenta que los tiempos de transmisión actuales sólo permiten la visualización cómoda de documentos con pocos Kbytes de gráficos, así un gráfico (o un conjunto de) cuyo tamaño no supere lo 30 Kbytes se considera óptimo, arriba de 30 y hasta 60 kbytes es bueno aunque un poco lento; más allá de 60 Kbytes no es recomendable.

*Apêndice A*

## Apéndice B

### ***Organismos Internacionales y Proveedores de acceso en México***

---

Conforme Internet creció surgió la necesidad de asegurar su desarrollo y evolución en forma ordenada, de manera que ésta no llegara a convertirse en un "caos". Para tal propósito fueron creados varios organismos, que al ir evolucionando conformaron lo que hoy se conoce como Internet Society (ISOC).

Sus principales objetivos son:

- Desarrollo, mantenimiento, difusión y evolución de las normas en Internet.
- Crecimiento y evolución de la arquitectura.
- Asistencia técnica a los países en desarrollo.
- Colección y difusión de información acerca de Internet
- Coordinación con otros gobiernos e instituciones para lograr la efectividad de los anteriores propósitos.

Los miembros de este organismo están obligados por común acuerdo al mantenimiento, viabilidad y acceso global de Internet, entre ellos destacan compañías, agencias de gobierno y fundaciones que han sido creadas para Internet.

ISOC se encuentra dividida en varias organizaciones que regulan actividades específicas de la red global, éstas son:

#### ***The Internet Architecture Board (IAB)***

La cual tiene a su cargo lo referente a propuestas técnicas y políticas que envuelven la evolución de la arquitectura Internet. Entre sus funciones se encuentran: revisar los estándares, actuar como regulador internacional de políticas técnicas entre la comunidad Internet.

#### ***The Internet Engineering Task Force (IETF)***

Es un grupo de diseñadores de redes, operadores, e investigadores relacionados con Internet y sus protocolos, que ayudan en la operación, manejo y evolución de la red global.

***The Internet Research Task Force (IRTF)***

Tiene como principal función motivar la investigación en interconexión de redes y desarrollos de nuevas tecnologías, la cual esta compuesta por grupos de investigación.

***The Network Information Center (NIC)***

Tiene como función la de proveer información acerca de servicios que ofrece la red.

***InterNIC***

Es un proyecto donde colaboran tres organizaciones: General Atomics que provee servicios de información desde sus instalaciones en San Diego, CA., AT&T que suministra servicios de bases de datos desde South Plainsfield, NJ; y Network Solutions, Inc. proporcionando servicios de registro, tales como reparto de direcciones IP, registro de dominio y asignación de ASN (Autonomous System Number assignment), desde Herndon, VA.

***The Internet Registry (IR)***

Es la responsable de asignar identificadores, tales como números de red IP, sistemas autónomos de números para redes, etc.

## Proveedores de acceso en México:

La información contenida en el presente directorio NO es de lo más confiable posible, ya que en un mercado con un ritmo tan rápido, resulta muy difícil garantizar la permanencia de algunos proveedores de servicios. Probablemente para cuando Usted lea las siguientes líneas, hayan nacido nuevas compañías y existe la posibilidad de que algunos de los proveedores listados, presenten cambios o inclusive hayan desaparecido:

### **ALBEC Telecomunicaciones**

*Albec, S.A. de C.V.*  
 e-mail: [info@albec.net.mx](mailto:info@albec.net.mx)  
<http://www.albec.net.mx>

### **Baja OnLine**

*Multiservicios Geminis, S.A. de C.V.*  
 e-mail: [sysop@onyx.cachanilla.com.mx](mailto:sysop@onyx.cachanilla.com.mx)  
<http://www.cachanilla.com.mx>

### **BSM**

*Brain systems México, S.A. de C.V.*  
 e-mail: [vtas@www.bsmx.com](mailto:vtas@www.bsmx.com)  
<http://www.bsmx.com>

### **CIATEQ**

*CIATEQ, A.C.*  
 e-mail: [webmaster@ciateq.mx](mailto:webmaster@ciateq.mx)  
<http://www.ciateq.mx>

### **DataNet**

*Datanet, S.A. de C.V.*  
 e-mail: [info@data.net.mx](mailto:info@data.net.mx)  
<http://www.data.net.mx>

### **Ges-Internet**

*Global Enterprise Services de México, S.A.*  
 e-mail: [info@mx.ges.com](mailto:info@mx.ges.com)  
<http://www.mx.ges.com>

### **IBM Global Networks**

*IBM de México, S.A. de C.V.*  
 e-mail: [mexign@mexvm2.vnet.ibm.com](mailto:mexign@mexvm2.vnet.ibm.com)  
<http://www.ibm.com>

### **Infored de México**

*Info Red de México, S.A. de C.V.*  
 e-mail: [info@mail.inforedmx.com.mx](mailto:info@mail.inforedmx.com.mx)  
<http://www.inforedmx.com.mx>

### **Antequera**

*Antequera Red, S.A. de C.V.*  
 e-mail: [pepe@antequera.com](mailto:pepe@antequera.com)  
<http://antequera.com>

### **Bravo Net**

*Internet de Nuevo Laredo, S.A. de C.V.*  
 e-mail: [milton@rio.bravo.net](mailto:milton@rio.bravo.net)  
<http://www.bravo.net>

### **1CaboNet**

*Redes de los Cabos, S.A. de C.V.*  
 e-mail: [mklein@1cabonet.com.mx](mailto:mklein@1cabonet.com.mx)  
<http://www.1cabonet.com.mx>

### **CompuServe**

*Infoaccess, S.A. de C.V.*  
 e-mail: [7006.101@compuserve.com](mailto:7006.101@compuserve.com)  
<http://www.compuserve.com.mx>

### **Dyred**

*Enlaces y Comunicaciones del Sureste*  
 e-mail: [dyred@sureste.com](mailto:dyred@sureste.com)  
<http://www.dyred.sureste.com>

### **Gigacom**

*Gigacom, S.A. de C.V.*  
 e-mail: [info@giga.com](mailto:info@giga.com)  
<http://www.giga.com>

### **Infolink**

*Infolink, S.A. de C.V.*  
 e-mail: [informes@infolink.net.mx](mailto:informes@infolink.net.mx)  
<http://www.infolink.net.mx>

### **Infosel**

*Información Selectiva, S.A. de C.V.*  
 e-mail: [info@infosel.com.mx](mailto:info@infosel.com.mx)  
<http://www.infosel.com.mx>

**Intermex**

*Net-Intermex, S.A. de C.V.*  
e-mail: [info@mail.intermex.com.mx](mailto:info@mail.intermex.com.mx)  
<http://www.intermex.com.mx>

**México en Red**

*Asesoría en Sistemas Automatizados para Empresas, S.A. de C.V.*  
e-mail: [huibert@mail.proesa.com.mx](mailto:huibert@mail.proesa.com.mx)  
<http://www.mexred.com.mx>

**MPSNet**

*Direct Tech, S.A. de C.V.*  
e-mail: [webmaster@mpsnet.com.mx](mailto:webmaster@mpsnet.com.mx)  
<http://www.mpsnet.com.mx>

**Planet**

*Interplanet, S.A. de C.V.*  
e-mail: [info@planet.com.mx](mailto:info@planet.com.mx)  
<http://www.planet.com.mx>

**Pvnet**

*Pvnet, S.A. de C.V. (Puerto Vallarta)*  
e-mail: [info@pvnet.com.mx](mailto:info@pvnet.com.mx)  
<http://www.pvnet.com.mx>

**RTI Net**

*Redes y Tecnología de Información, S.A.*  
e-mail: [info@rti.net.mx](mailto:info@rti.net.mx)  
<http://www.rti.net>

**SPIN**

*Tecnología Uno-Cero, S.A. de C.V.*  
e-mail: [ventas@spin.com.mx](mailto:ventas@spin.com.mx)  
<http://www.spin.com.mx>

**Logicnet**

*Compulogic, S.A. de C.V.*  
e-mail: [anavarro@gauss.logicnet.com.mx](mailto:anavarro@gauss.logicnet.com.mx)  
<http://www.logicnet.com.mx>

**Mundo Internet**

*CINVESTAV-IPN, Unida Mérida*  
e-mail:  
[webmaster@minter.cieamer.conacyt.mx](mailto:webmaster@minter.cieamer.conacyt.mx)  
<http://minter.cieamer.conacyt.mx>

**Networks México**

*Networks México, S.A. de C.V.*  
e-mail: [info@netmex.com](mailto:info@netmex.com)  
<http://www.netmex.com>

**Proveedor de acceso corporativo**

*ITESM, Campus Edo. de México*  
e-mail: [tecomp@campus.cem.itesm.mx](mailto:tecomp@campus.cem.itesm.mx)  
<http://www.cem.itesm.mx>

**Red Tecnológica Nacional**

*Infotec, S.A.*  
e-mail: [infotec@rtn.net.mx](mailto:infotec@rtn.net.mx)  
<http://www.rtn.net.com>

**Soluciones avanzadas de redes**

*Soluciones avanzadas de redes, S.A.*  
e-mail: [info@sar.net](mailto:info@sar.net)  
<http://www.sar.net>

**SuperNet**

*Supernet, S.A. de C.V.*  
e-mail: [info@supernet.com.mx](mailto:info@supernet.com.mx)  
<http://www.supernet.com.mx>

# Apéndice C

## **Estándares de Comunicación**

---

Se pueden dividir las organizaciones de estándares, en dos grupos principales, uno donde se encuentra la ANSI (American National Standards Institute) y el IEEE (Institute of Electrical and Electronics Engineers) en los EUA y otro que agrupa a organismos internacionales como CCITT (Consultative Committee of International Telephone and Telegraph) y la ISO (International Standard Organization). Las organizaciones estadounidenses han desarrollado principalmente estándares para redes de área local, mientras que la organizaciones internacionales han desarrollado numerosos estándares para facilitar la operación de redes de área local y área amplia.

La ISO delegó la normalización de redes de área local a ANSI, la cual a su vez delegó a la IEEE los estándares para LAN de velocidades inferiores a 50 Mbps. Con esto ANSI ha desarrollado el estándar para el FDDI (Fiber Distributed Data Interface) que opera a 100 Mbps, en tanto que la IEEE los desarrolló los estándares para Ethernet, Token Ring, y otros tipos de redes de área local.

### **ANSI**

Es la principal organización de estándares en los Estados Unidos de Norteamérica, institución de carácter no lucrativo que representa a este país ante la ISO. Los estándares ANSI se desarrollan a través del trabajo de aproximadamente 300 comités de estándares, y con la participación de asociaciones como la EIA (Electronic Industry Association). Para la industria de la computación se estableció el Comité de Estándares X en 1960, éste firmado a su vez por 25 comités técnicos, cada uno asignado a desarrollar un estándar para una área técnica específica. Uno de ellos es el Comité X3S3, formalmente conocido como el Comité Técnico de Comunicaciones de Datos, el cual desarrolló el estándar ANSI X3T9.5, el cual controla la operación de FDDI, conocido en la actualidad como estándar ISO 9314.

### **IEEE**

El Instituto de Ingenieros Eléctricos y Electrónicos (IEEE), es una sociedad de ingenieros, la cual ha sido muy activa en el desarrollo de estándares de comunicaciones de datos. De

hecho, se trata de la más prominente desarrolladora de estándares para redes locales, cuyo subcomité 802 comenzó a trabajar en 1980 antes de que se estableciera un mercado estable y viable para estas tecnologías. El proyecto IEEE 802, concentra sus esfuerzos en la normalización de las interfaces físicas, entre los dispositivos de la red y los procedimientos y funciones que se requieren para establecer, mantener y terminar la conexión entre ellos. Estos procedimientos incluyen la definición del formato de los datos, procedimientos de control de error y otras actividades de control que normen el flujo de información. Este enfoque de los estándares de la IEEE hace que actualmente representen los dos niveles inferiores del modelo OSI (físico y de enlace).

## **CCITT**

El Comité Consultivo Internacional de Telefonía y Telegrafía (CCITT) es un grupo entre la Unión Internacional de Telecomunicaciones, la cual es una agencia especializada de la Naciones Unidas. La CCITT tiene la responsabilidad directa del desarrollo de estándares de comunicación de datos, compuesta por 15 grupos de estudio, cada uno de los cuales tiene un área específica de responsabilidad. El trabajo de la CCITT está desarrollado en ciclos de cuatro años, conocido como Período de Estudio. Al terminar cada período, se lleva a cabo una sesión plenaria, para dar revisión a los cuatro años pasados, recomendaciones propuestas son consideradas para su adopción, y los temas a ser investigados durante el siguiente período son considerados.

La décima sesión plenaria del CCITT fue en 1992, aunque las recomendaciones aprobadas no tienen carácter obligatorio, éstas tienen efectos en las leyes de algunos países europeos, muchas de estas recomendaciones han sido adoptadas por proveedores de servicios de comunicación en los EUA. El juego de recomendaciones CCITT más conocido son las series V, las cuales describen la operación de diferentes características en módem, como puede ser la compresión de datos, detección y corrección de errores en la transmisión de los mismos.

La CCITT, ahora es mejor conocida como ITU-T, Sector de Estandarización en Telecomunicaciones de la Unión Internacional de Telecomunicaciones.

## **ISO**

La Organización de Estándares Internacional (ISO), es una entidad no gubernamental que tiene un carácter de consultor en el Consejo Económico Social de las Naciones Unidas. Su objetivo es promover el desarrollo de estándares en el mundo, con la finalidad de facilitar el intercambio internacional de bienes y servicios.

Los miembros del ISO consisten en las organizaciones nacionales de estándares de la mayoría de los países. Existen aproximadamente 100 países con participación directa en este organismo. El más notable estándar propuesto por ISO en el campo de las comunicaciones, es el desarrollo del modelo de referencia para interconexión de sistemas abiertos (OSI: Open Systems Interconnection), el cual contiene siete niveles:

Nivel 7	Capa de Aplicación
Nivel 6	Capa de Presentación
Nivel 5	Capa de Sesión
Nivel 4	Capa de Transporte
Nivel 3	Capa de Red
Nivel 2	Capa de Enlace de Datos
Nivel 1	Capa Física

**Modelo de Referencia OSI**

**Modelo OSI**

*(Open System Interconnection)*

Para sistemas abiertos se emplean protocolos estratificados por las siguientes razones: los organismos de normalización más importantes como ISO, CCITT, ECMA y ANSI entre otros, apoyaron la utilización de estas, además los principales fabricantes han instalado protocolos estratificados. Aunque la principal razón es haberse basado en el sentido común y estar respaldados técnicamente, para descomponer de manera lógica una red en partes (niveles o estratos) más pequeñas y fáciles de entender.

Lo anterior proporciona interfaces normalizadas entre distintas funciones de la red, conseguir simetrías en las funciones que se realizan en cada nodo, ofrecer un método que permita predecir y controlar posibles cambios en la lógica de la red y por último establecer un lenguaje normalizado que pueda clarificar la comunicación entre los distintos diseñadores, fabricantes, distribuidores y usuarios de redes.

De esta manera ISO y CCITT crearon un modelo para interconexión de sistemas abiertos llamado Open System Interconnection (OSI) cuyos objetivos principales son:

- Proporcionar normas para la comunicación de sistemas.
- Abstracta el funcionamiento interno de los sistemas individuales.
- Eliminar todos los impedimentos técnicos que pudieran existir en la comunicación de sistemas.
- Definir los puntos de interconexión para el intercambio de información en sistemas.
- Ofrecer un punto de partida válido, a partir de comenzar en caso de que las definiciones del estándar no satisfagan todas las necesidades.
- Limitar el número de opciones, para incrementar las posibilidades de comunicación sin necesidad de muchas conversiones y traducciones a diferentes productos.

Este Modelo está formado por siete niveles, cada uno de los cuales especifica funciones particulares para la red. El nivel más alto es el cercano al usuario y el nivel más bajo es el cercano a la tecnología del medio físico.

El modelo OSI es universalmente usado como método de enseñar y entender la funcionalidad de las redes.



<b>ANSI</b>	American National Standards Institute. Organismo de los EUA, encargado de establecer los estándares de telecomunicaciones en ese país.
<b>ancho de banda</b>	Capacidad máxima de transmisión de un enlace. Usualmente se mide en bits por segundo (bps).
<b>API</b>	Application Programming Interface. Interfaz estándar utilizada en el desarrollo de programas fuentes de aplicaciones, que asegura la compatibilidad de las aplicaciones en todas las plataformas que soportan API, incluyendo plataformas de diferentes arquitecturas. Esto significa que sólo una versión de una aplicación en API necesita ser generada, y posteriormente el desarrollador sólo recompila el programa fuente en cada una de las plataformas en la cual la aplicación va a ejecutarse.
<b>aplicación</b>	Programa que realiza un conjunto particular de funciones relacionadas. Las aplicaciones pueden ser horizontales (común a varios departamentos) como procesadores de palabras, hojas de cálculo, bases de datos, correo electrónico, o verticales (específicas según sea el caso), como modelación médica, facturación, algoritmos de visualización.
<b>archivo binario</b>	También llamado binario, es un archivo que contiene código de bajo nivel que puede ser leído y ejecutado por el sistema operativo. Un binario es creado cuando el código fuente, es compilado (no todos los binarios son ejecutables).
<b>ARP</b>	Address Resolution Protocol. Protocolo de Resolución de Direcciones. Se encarga de proporcionar dinámicamente el mapeo entre una dirección internet y la dirección física de hardware.
<b>AT&amp;T</b>	AT&T Bell Labs. Es el equipo de trabajo, investigación y desarrollo de AT&T, responsable del desarrollo original y evolución de Unix. La responsabilidad de Unix fue transferida a USL (Unix System Labs), un derivado de AT&T actualmente subsidiaria de Novell.
<b>background</b>	Proceso que no interactúa de manera visible con el usuario, el cual se ejecuta comúnmente en un entorno multitareas, asignándole ciertas prioridades de procesamiento.  En WWW, fondo base de la aplicación Web, es el color de la pantalla en donde se aloja el texto, imágenes, ligas, etc.
<b>boot</b>	Es el conjunto de procesos mediante los cuales se inicializan programas de arranque de un sistema de cómputo, a través de la ejecución automática de instrucciones necesarias para tal fin.

## Glosario

<b>cache</b>	Un cache son datos que guarda el sistema operativo o diversos subsistemas de entrada/salida, que no son requeridos en un momento determinado, pero pueden ser de gran utilidad en el momento de realizar una nueva operación como lectura o recuperación de información.
<b>cliente</b>	Sistema o programa de computadora que realiza peticiones de servicio a otro programa o sistema (comúnmente llamado servidor).
<b>cliente/ servidor</b>	Modelo de trabajo basado en red, el proceso de manejo de información se encuentra distribuido entre un proceso cliente y un proceso servidor, usualmente en distintas máquinas o computadoras.
<b>compilador</b>	Programa que traduce el código fuente (lenguaje de alto nivel) en un código ensamblador, el cual luego es ligado para crear un archivo binario que puede ser ejecutado por el hardware de la computadora sin ningún problema.
<b>datagrama</b>	Mensaje enviado en una red de conmutación de paquetes.
<b>daemon</b>	Programa o aplicación que espera generalmente en modo background, listo para ejecutarse cuando se cumplan ciertas condiciones o determinado evento.
<b>DGSCA</b>	Dirección General de Servicios de Cómputo Académico (UNAM).
<b>DNS</b>	Domain Name Service. Servicio de Dominio de Nombres, base de datos distribuida que traduce nombres de computadoras a sus correspondientes direcciones IP.
<b>Ethernet</b>	Protocolo de comunicación (conexión) ampliamente utilizado en redes de área local, desarrollado por Xerox Corp. Los datos se transmiten a 10 megabits por segundo. El estándar ethernet está definido por la especificación IEEE 802.3
<b>FAQ</b>	Frequently Asked Question. Preguntas Más Frecuentes, es un conjunto de archivos que contienen preguntas comunes al tema específico al que haga referencia así como sus respuestas.
<b>FDDI</b>	Fiber Data Distributed Interface. Protocolo (parecido al token ring) que utiliza fibra óptica como medio de transmisión.
<b>fibra óptica</b>	Medio de transmisión de datos por el cual viajan pequeñas cantidades de luz producidas por un rayo láser o un diodo emisor de luz (led).
<b>firewall</b>	Mecanismo utilizado para proteger una red (como antiguamente "pared contra incendios", para proteger construcciones de madera), o computadora conectada a Internet de accesos no autorizados. Un firewall, puede implementarse con software, hardware o con una combinación de ambos.
<b>FTP</b>	File Transfer Protocol. Protocolo de Transferencia de Archivos, este define los mecanismos y reglas para transferir copias de archivos entre las diversas computadoras de la "red de redes".
<b>freeware</b>	Software que ha sido puesto a disposición de la comunidad por sus autores. Este tipo de programas son libremente distribuidos y utilizados sin necesidad de pago alguno.

<b>gateway</b>	Dispositivo que interconecta y realiza la conversión de protocolos entre dos tipos de redes o aplicaciones, se trata de todo aquel sistema que permita el intercambio de información entre diferentes estándares de una aplicación.
<b>GIF</b>	Graphics Interchange Format. Formato de Intercambio de Gráficos, es un formato para imágenes gráficas, estándar facto en Internet.
<b>Gopher</b>	Sistema de navegación en Internet basado en el modelo cliente/servidor, a través de menús. Puede verse como un precursor del Web y, aunque paulatinamente se encuentra desapareciendo, aún quedan miles de servidores Gopher en servicio.
<b>GUI</b>	Graphical User Interface. Interfaz Gráfica al Usuario.
<b>hipermedia</b>	Extensión del concepto de hipertexto para la inclusión del manejo de información en forma de imágenes, gráficas, sonido o video.
<b>hipertexto</b>	Termino usado para referirse a un sistema no lineal de búsqueda y recuperación de información que actúa mediante hiperligas. Usualmente un documento Web contiene enlaces dentro y/o hacia otros documentos, estos enlaces son llamados hiperligas o simplemente ligas.
<b>hiperliga</b>	Código de un documento HTML, el cual permite efectuar "brincos" o "saltos" hacia otro lugar, ya sea, en el mismo documento en el servidor local o algún otro en servidores remotos.
<b>host</b>	Computadora que trabaja dentro de una red, particularmente un transmisor o receptor desde el punto de vista de la red global de comunicación Internet.
<b>HTML</b>	HyperText Mark-up Language. Lenguaje de Marcación de Hipertexto, utilizado para dar formato a la información que es utilizada en WWW; por sí mismo es un derivado de SGML, que fue un estándar desarrollado a mediados de 1980.
<b>HTTP</b>	HyperText Transport Protocol. Protocolo de Transporte de Hipertexto. Se basa en la arquitectura cliente-servidor.
<b>Home page</b>	Se trata únicamente de la página principal de un sitio Web (Web site)
<b>interfaz</b>	Medio que sirve de enlace entre las partes de un sistema de transmisión.
<b>Internet</b>	En minúsculas, cualquier colección de redes distintas trabajando juntas como una misma.  Internet (especialmente en l mayúscula), la red global de redes conectadas a través del conjunto de protocolos TCP/IP.
<b>Intranet</b>	Red de uso privado (de cualquier magnitud) que emplea los mismos estándares, aplicaciones y herramientas de Internet, a través de TCP/IP.
<b>IO</b>	Input Output. Entrada/salida.
<b>IP</b>	Internet Protocol. Este protocolo permite que paquetes de información atraviesen múltiples redes para alcanzar su destino final.

<b>ISO</b>	International Organization for Standardization (Organización Internacional de Estándares).
<b>ISOC</b>	The Internet Society (Sociedad Internet).
<b>JAVA</b>	Lenguaje de programación orientado a objetos, así como plataforma de hardware (independiente del lenguaje), creado por Sun Microsystems. Diseñado para una arquitectura cliente-servidor, en la que sólo es necesario el intercambio de pequeñas porciones de código (llamadas Applets), que son ejecutadas por el cliente.
<b>JavaScript</b>	Lenguaje de programación basado en objetos creado por Netscape, para trasladar ciertas tareas simples del lado del cliente, procurando acceder al mínimo el servidor Web.
<b>LAN</b>	Local Area Network (Red de Area Local).
<b>login</b>	Se refiere al nombre del usuario o identificador del mismo en la computadora.
<b>NNTP</b>	Network News Transfer Protocol. Protocolo de Transferencia de Noticias en Red.
<b>MAN</b>	Metropolitan Area Network. Red de Area Metropolitana.
<b>MIME</b>	Multi-purpose Internet Mail Extensions. Extensiones de Correo Internet Multipropósito. Extensión al protocolo de correo de Internet que permite el intercambio de archivos binarios como anexos de mensajes de correo electrónico. MIME también es empleado por servidores Web para identificar el tipo de archivos que envían a los clientes (visualizadores).
<b>Mosaic</b>	Visualizador de WWW que puede ejecutarse en diversas plataformas, fué el primer cliente para servidores Web.
<b>navegar</b>	Término empleado cuando se "viaja" entre documentos hipertexto en World Wide Web, a través de los diversas ligas contenidas en el documento (suele también emplearse el término "surfear").
<b>NCSA</b>	National Center for Supercomputing Applications, (creador de Mosaic).
<b>NIC</b>	Network Information Center. Centro de Información de Red, responsable de coordinar las funciones de esta, así como proporcionar información.
<b>NIS</b>	Network Information System. Sistema de Información de Red.
<b>NFS</b>	Network File System. Sistema de Archivos de Red.
<b>NOC</b>	Network Operation Center. Centro de Operación de Red, grupo responsable del cuidado diario de una red.
<b>online</b>	Término que puede ser traducido como "conexión activa", es cuando el visualizador se encuentra con una conexión real (temporal), y disponible para efectuar enlaces a cualquier sitio con un URL definido.
<b>password</b>	Se refiere a la clave del usuario para poder acceder a la computadora.

<b>PERL</b>	Practical Extraction and Reporting Language. Lenguaje útil para fines de programación de CGI's en un servidor WWW, comúnmente utilizado en ambientes UNIX.
<b>PPP</b>	Point to Point Protocol; Protocolo Punto a Punto. Es empleado comúnmente para efectuar conexiones TCP/IP a través de enlaces seriales.
<b>protocolo</b>	Conjunto de reglas que gobiernan el formato y significado de los paquetes o mensajes, que se utilizan al establecer una comunicación entre dos o más computadoras.
<b>puerto</b>	Específicamente en Internet, se refiere a la parte de un URL que va inmediatamente después de un nombre de dominio y que se encuentra precedido por dos puntos (:). Se emplea para indicar que los servicios de dicho servidor no están ejecutándose en el puerto estándar.
<b>RARP</b>	Reverse Address Resolution Protocol. Protocolo de Resolución de Nombres Inverso.
<b>RFC</b>	Requests For Comments, serie de escritos y documentos que establecen los estándares y propuestas sobre Internet. Cuando se propone un nuevo estándar, la IETF publica un RFC, el cual a su vez le hereda el nombre al estándar finalmente adoptado.
<b>ruteador</b>	Dispositivo que provee enlace entre una red y otra.
<b>servidor</b>	Sistema que permite a otro (comúnmente llamado cliente) interactuar con él, para permitirle que lleve a cabo diversas tareas o procesos.
<b>script</b>	Colección de comandos nativos de un sistema operativo almacenados en un archivo, que son ejecutados secuencialmente por este.
<b>shareware</b>	Software de distribución pública, pero no de uso gratuito. El autor o distribuidor, establece un período de prueba después del cual solicita una cuota de recuperación.
<b>SGML</b>	Standard Generalized Markup-Language. Lenguaje Estándar Generalizado de Marcación; estándar de ISO utilizado para marcar documentos.
<b>SLIP</b>	Serial Line Internet Protocol. Protocolo Internet de Línea Serial (a través de línea telefónica).
<b>SMTP</b>	Simple Mail Transfer Protocol. Protocolo Simple o sencillo de Transferencia de Correo. Sólo permite el intercambio de mensajes ASCII, por lo que esta siendo gradualmente reemplazado por MIME.
<b>socket</b>	Componente lógico de los protocolos TCP/IP, que actúa como punto inicial y final de la comunicación de procesos entre computadoras distintas.
<b>TCP/IP</b>	Transmission Control Protocol/Internet Protocol. Protocolo de Control de Transmisión/Protocolo Internet. Diseñados inicialmente para ambientes UNIX, por Vinton G. Cerf y Robert E. Kahn, son empleados para la comunicación de todos los dispositivos involucrados en la "red de redes".
<b>Telnet</b>	Protocolo de emulación de terminal, mediante el cual un usuario común puede entrar a sesión en sistemas remotos.

## *Glosario*

<b>UDP</b>	User Datagram Protocol. Protocolo de Datagrama de Usuario.
<b>UNIX</b>	Sistema operativo, el término también es aplicable a la clase de computadoras que pueden ejecutar este S.O.
<b>URL</b>	Uniform Resource Locator. Localizador Uniforme de Recursos.
<b>visualizador</b>	Programa de aplicación mediante el cual se pueden observar gráficamente documentos enlazados en todo el mundo a través de la red WWW, es un medio de acceso a Internet.
<b>WAN</b>	Wide Area Network. Red de Area Amplia, comúnmente compuesta por la conexión de varias LAN.
<b>WAIS</b>	Wide Area Information Server. Servidor de Información de Area Amplia, servicio de Internet utilizado para localizar información principalmente en bases de datos.
<b>Web page</b>	Página de Web, cualquier página en un sitio Web o Web site.
<b>Web site</b>	Sitio Web, conjunto de páginas Web que forman una unidad única. Incluso se puede tener un sitio Web de una sola página.
<b>WYSIWYG</b>	What You See is What You Get. "Lo que ves es lo que obtienes", comúnmente empleado en programas de aplicación como: diseño de documentos de calidad o páginas de WWW.

# Referencias

---

## Bibliografía:

Chandler  
***Running a Perfect Web Site***  
1995 QUE Corporation  
ISBN: 0-7897-0210-X

Commer, Douglas  
***Internetworking with TCP/IP, vol. 1***  
1995 Prentice Hall

Dick Oliver  
***Netscape 2 Unleashed***  
1996 Sams.net Publishing  
ISBN: 1-57521-007-X

Yeager, McGrath  
***Web Server Technology***  
1996 Krol Publishing  
ISBN: 1-55860-376-X

***The Internet Catalog & User's Guide***  
1992 O'Reilly & Associates Inc.

***Using Netscape***  
1995 QUE Corporation  
ISBN: 1-57521-007-X

Cricket, Jerry  
***Managing Internet Information Services***  
1994 O'Reilly & Associates Inc.

December, Ginsburg  
***HTML & CGI Unleashed***  
1995 Sams.net Publishing  
ISBN: 0-672-30745-6

Fah-Chun Cheong  
***Internet Agents***  
1996 New Riders Publishing  
ISBN: 1-56205-463-5

Fox, Downing  
***HTML Web Publisher's Construction Kit***  
1996 White Group Press  
ISBN: 1-57169-018-2

***JAVA Unleashed***  
1996 Sams.net Publishing  
ISBN: 1-57521-049-5

- Cruz Cortés César

***"Uso de HTML y CGI para el desarrollo de aplicaciones en Internet"***

1997, ENEP Aragón, UNAM.  
Director de Tesis: Ing. Martín Ordoñez Rosales  
ICOM, (Tesis en desarrollo).

- Mejía Berzunza Ivonne

***"Panorámica general de los principios y elementos básicos que intervienen en la super carretera de información"***

1996, ENEP Aragón, UNAM.  
Director de Tesis: Ing. Ernesto Peñaloza R.  
ICOM, Tesis No. 282

- Mendieta M. Marisela

***"Descripción y funcionamiento de MOSAIC, herramienta para navegar Internet"***

1995, ENEP Aragón, UNAM.  
Director de Tesis: Ing. Silvia Vega M.  
ICOM, Tesis No. 210

- Gómez Luviano Antonio

***"TCP/IP conceptos y aplicaciones"***

1994, ENEP Aragón, UNAM.  
ICOM, Tesis No. 177

**URLs:**

**WWW**

- The WWW Consortium (W3C)  
<http://www.w3.org/pub/www>
- Internet Society Home Page  
<http://www.isoc.org/home.html>
- What is the Internet Society?  
<http://www.isoc.org/whatis>
- A History of Internet Protocols at CERN  
(Ben M. Segal)  
<http://wwwcn.cern.ch/pdp/ns/ben/TCPHIST.html>
- How to Publish on the Web  
<http://www.thegiim.org/>
- Design and Style Guidelines  
<http://WWW.Stars.com/Tutorial/Style/>
- Sun Guide to Web Style  
<http://www.sun.com/styleguide/>
- The Web Designer  
<http://web.canlink.com/webdesign/>
- The Common Gateway Interface (CGI)  
<http://hooohoo.ncsa.uiuc.edu/cgi/>
- Perl  
<http://www.perl.com/perl/>
- NCSA HTTPd  
<http://hooohoo.ncsa.uiuc.edu/docs/Overview.html>  
<http://space.acm.ndsu.nodak.edu/~fgonzale/terms.html>

**FTP**

<ftp://ftp.cdrom.com/pub/perl/CPAN/doc/manual/html/perl.html>