

107  
zej



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

FACULTAD DE CIENCIAS

UNA METODOLOGIA PARA LA CREACION DE  
SITIOS WEB Y SU APLICACION

**T E S I S**

QUE PARA OBTENER EL TITULO DE:

**A C T U A R I O**

**P R E S E N T A :**

**ALEJANDRO TALAVERA ROSALES**



DIRECTOR DE TESIS: M. EN I.O. VICTOR RAFAEL PEREZ PEREZ

1997

**TESIS CON  
FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

PAGINACION VARIA

COMPLETA LA INFORMACION



UNIVERSIDAD NACIONAL  
AVENIDA DE  
MÉXICO

M. en C. Virginia Abrín Batule  
Jefe de la División de Estudios Profesionales de la  
Facultad de Ciencias  
P r e s e n t e

Comunicamos a usted que hemos revisado el trabajo de Tesis:

Una metodología para la creación de sitios Web y su aplicación.

realizado por Alejandro Lalavera Rosales

con número de cuenta RR28129-4 , pasante de la carrera de Actuaría.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis  
Propietario

M. en I.O. Víctor Rafael Pérez Pérez

Propietario

M. en C. María Guadalupe Elena Ibarquiengoitia González

Propietario

M. en I.O. María de Luz Gasca Soto

Suplente

Act. Bárbara Ruth Trejo Becerril

Suplente

Act. Javier Terán Castro

*Barbara Trejo*  
TERÁN  
CASTRO

*que es la guardadora*

Consejo Departamental de Matemáticas

Act. Agustín Román Aguilar

Doy gracias a todas las personas que por su apoyo y confianza me han ayudado a dar un paso más en el camino de la vida, y a todas ellas les dedico este trabajo :

A mis padres José Luis y Silvia, que con su guía y atenciones me han formado.

A mis hermanos Marco Antonio y Artemisa, que siempre estuvieron apoyándome.

A mis amigos Bárbara, Carlos, Javier, Juan Carlos, Laura, Leticia, Luz, Marissa, Miguel Ángel, Mónica, Rodrigo, Rosalía y Thelma, por su apoyo incondicional que me han mostrado en más de una ocasión.

A mis maestros.

A Víctor R. Pérez P., mi director de tesis.

A mis sinodales.

A la U.N.A.M.

# ÍNDICE

## Índice

Introducción.....	3
1. Conceptos básicos.....	7
1.1. Redes de Computadoras.....	7
1.2. Clasificación de redes.....	8
1.3. Protocolos de comunicación.....	12
1.4. Los protocolos TCP/IP.....	14
1.5. La Internet.....	16
1.6. Tipos de acceso a Internet.....	17
2. Servicios en Internet.....	21
2.1. Internet y el modelo cliente/servidor.....	21
2.2. Telnet.....	23
2.3. Correo electrónico.....	28
2.4. Transferencia de archivos - FTP ( <i>File Transport Protocol</i> ).....	32
2.5. Búsqueda y recuperación de información.....	34
2.6. WWW (World Wide Web).....	35
3. Metodología para el desarrollo de sitios WWW.....	45
3.1. Presentación de la metodología.....	45
3.2. Elementos en el desarrollo de sitios WWW.....	46
3.3. Procesos del desarrollo.....	47
3.3.1. Planeación.....	49
3.3.2. Análisis.....	53
3.3.3. Diseño.....	60
3.3.4. Implementación.....	66
3.3.5. Promoción.....	68
3.3.6. Innovación.....	69
4. Aplicación de la metodología.....	73
4.1. Antecedentes.....	73
4.2. Planeación.....	74
4.3. Análisis.....	75
4.4. Diseño.....	76
4.5. Implementación.....	80
4.6. Promoción.....	86
4.7. Innovación.....	86
Conclusiones.....	93
Anexo A - El lenguaje HTML	
Anexo B - CGI (Common Gateway Interface)	
Glosario de términos de Internet	
Bibliografía	

# INTRODUCCIÓN

## Introducción

La distribución de tareas entre los grupos humanos es una solución natural a las grandes cargas de trabajo a las cuales se han enfrentado desde tiempos inmemoriales. Este tipo de técnicas han creado grupos de trabajo altamente especializados en cada una de las áreas del conocimiento. Un ejemplo de lo anterior lo podemos apreciar en las estructuras de organización de empresas públicas y privadas.

Si bien dicha distribución de trabajo permite planificar actividades que redunden en el beneficio propio de los grupos, ello se ve más enriquecido con el uso de equipo de cómputo. Así, las redes de computadoras han facilitado el establecimiento de grupos de trabajo remotos, con lo cual se pueden distribuir las cargas de trabajo en uno o más grupos de personas inclusive a nivel mundial.

Actualmente, la introducción de las redes y sus aplicaciones en muy diversas actividades, que no se limita solo a lo académico o administrativo, permiten un avance preciso, encaminado a solucionar problemáticas particulares. Así, es posible la explotación de las redes de computadoras para la distribución de programas, el establecimiento de negocios, la revisión de materiales de consulta, la divulgación de ideas y trabajos, entre muchos otros, todo esto, debido a los grandes avances que se han tenido en áreas tales como el cómputo y las telecomunicaciones.

Aunque existen muchos tipos de redes, el presente trabajo se centrará en el caso de la red *Internet*, debido al auge e importancia que ha presentado en años recientes. Mucho de su éxito se debe a la presentación de la información en una forma innovadora: el hipertexto. Este permite la asociación de documentos por medio de palabras claves, de tal manera que solo basta dar un clic con el *Mouse*, para que aparezca un nuevo documento. Esta forma de utilizar la red *Internet* se conoce como *World Wide Web (WWW)*.

Si bien, el crear documentos en la red para aprovechar la idea del hipertexto es relativamente sencillo, no lo es tanto cuando se pretende crear un conjunto de estos documentos para presentar un producto funcional que sea a la vez dinámico y atractivo para quienes lo consultan. Por tal motivo el presente trabajo tiene como principal objetivo, el proponer una metodología de desarrollo que permita la creación de sitios *WWW* para su puesta en funcionamiento, que satisfagan las necesidades del público que lo consulta.

## Introducción

El trabajo se desarrolla de la siguiente forma :

En el Capítulo 1 se presentan los conceptos generales acerca de las redes de computadoras necesarios para el desarrollo general del trabajo. Además, se muestra el funcionamiento particular del protocolo TCP/IP y se presenta la red *Internet*.

En el Capítulo 2 se presentan los servicios más usados en la red *Internet*, mostrando que a pesar de sus diferencias particulares, todo servicio en dicha red se basa en el modelo Cliente/Servidor; en especial el servicio de WWW, que se explica ampliamente.

En el Capítulo 3 se presenta la metodología para el desarrollo de sitios WWW. En dicha presentación se dan a conocer los elementos y procesos que se deben manejar, así como las relaciones que existen entre ellos.

En el Capítulo 4 se presenta el uso de la metodología en un caso práctico, para crear el sitio WWW del Departamento de Evaluación de Aplicaciones.

Finalmente, se muestran las conclusiones obtenidas a partir de la realización de este trabajo y la Bibliografía utilizada para su desarrollo.



# CAPÍTULO 1

## Conceptos básicos

### 1. Conceptos básicos

La comunicación de datos se ha convertido en un área de mucho interés en materia de cómputo. Gran parte de este interés se debe a las ventajas prácticas que proporciona la automatización de flujos de información en los diferentes sectores que conforman los núcleos de trabajo. Actualmente, existen diversas redes mundiales en las que podemos encontrar temas muy diversos, como las condiciones de los mercados de valores, noticias deportivas, resultados de encuestas, visitas a museos, servicios de suscripción a libros, etcétera. Un ejemplo de este tipo de redes es la red mundial de computadoras conocida como *Internet*.

Existen algunos servicios básicos que proporciona la red *Internet* que permiten a los usuarios comunicarse entre sí, y obtener información a través de ella.

El desarrollo de nuevos servicios utilizando *Internet* permiten resolver problemas de una forma muy eficaz. Por ejemplo, se pueden desarrollar: sistemas de inscripción, de consulta de calificaciones, o préstamos de libros, entre muchos otros. Con estos servicios, los beneficios se dan a todos niveles: mejor control en los flujos de información, ahorro considerable en el uso del papel, uso de la infraestructura que ya se tiene en los Campus Universitarios, etcétera.

En el presente capítulo se presentan conceptos y terminología sobre redes de computadoras, así como una breve descripción del desarrollo que la red *Internet* ha tenido, desde sus orígenes hasta los últimos años.

#### 1.1 Redes de Computadoras

Se denomina red de computadoras al conjunto de computadoras que se encuentran unidas por *hardware* y *software*, de tal manera que se les permiten compartir información y recursos.

Una de las razones por las cuales surgen las redes de computadoras, es la de distribuir recursos y la información que existen entre los equipos conectados.

La idea y el uso de redes de computadoras no es tan desconocida por las personas en general. Tomemos el caso de los cajeros bancarios. Dichos cajeros (computadoras), realizan transacciones de información con otros equipos; mediante ellos podemos consultar nuestro estado de cuenta, retirar efectivo, realizar pagos de servicios, entre otras operaciones. Cabe mencionar que para este ejemplo (y en general para cualquier red), nosotros somos usuarios finales.

## 1. Conceptos básicos

Como ya se ha definido, una red de computadoras consta de dos partes esenciales: una física (*Hardware*) y una lógica (*Software*). El *Hardware* permite la comunicación a través de dispositivos físicos, los cuales transmiten la información de una computadora a otra. El *Software*, por su parte, pone en funcionamiento dichos dispositivos así como el correcto manejo e interpretación de la información. Aunque estos son los elementos esenciales de una red, puede variar su tipo y funcionamiento de equipo a equipo.

### 1.2 Clasificación de Redes

Existen diversos tipos de redes, variando desde la forma en que están construidas, pasando por su alcance geográfico, hasta la forma de manejar la información. Para obtener una mejor visión de una red en particular, podemos clasificar a las redes de computadoras de acuerdo a su extensión, medio y topología.

#### Por su Extensión

Diferenciamos a las redes con esta clasificación por su extensión geográfica. De aquí podemos observar tres tipos de redes :

- LAN (*Local Area Network*) : Conocidas como redes de amplitud local por sus siglas en inglés. Dichas redes son "pequeñas", abarcando algunos cientos de metros. La Figura 1.1 ejemplifica una red LAN.

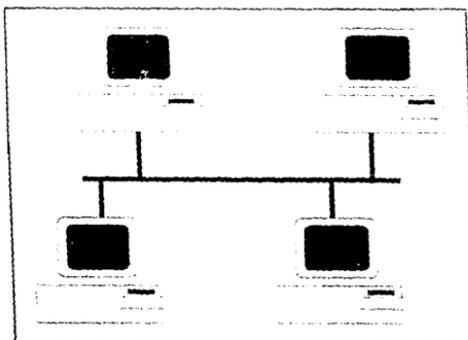


Figura 1.1 Red LAN

- **MAN (Metropolitan Area Network)** : Conocidas como redes de amplitud metropolitana por sus siglas en inglés. Dichas redes abarcan un área metropolitana; se componen por lo general de varias LAN's. La Figura 1.2 ejemplifica una red MAN.

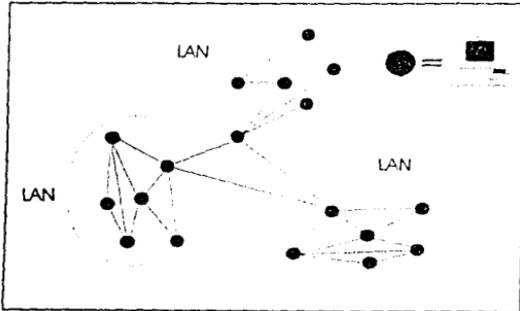


Figura 1. 2 Red MAN

- **WAN (World Area Network)** : Conocidas como redes de amplitud mundial por sus siglas en inglés. Dichas redes abarcan áreas a niveles internacionales; en general, se componen de varias MAN's. La Figura 1.3 ejemplifica una red WAN.

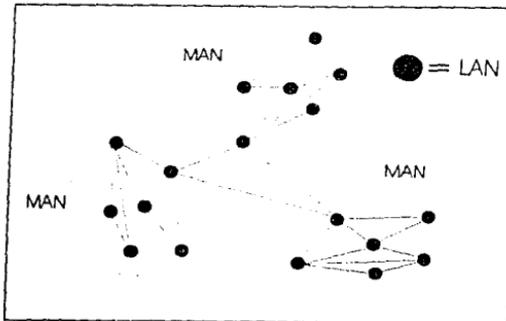


Figura 1. 3 Red WAN

## 1. Conceptos básicos

Siguiendo con el ejemplo de los cajeros automáticos, una red LAN puede ser un grupo de ellos en un centro comercial. Una red MAN estaría representada por el conjunto de cajeros de varios centros comerciales en la capital de la república (por citar una ciudad). Una red WAN sería aquella que se forma de varias redes de cajeros a nivel mundial.

Aunque esta clasificación no es precisa, debido a que es difícil establecer las magnitudes absolutas para cada tipo, nos proporciona una idea de su envergadura en general.

### Por el Medio

Esta clasificación se basa en el tipo de hardware que se utilizó en el diseño y construcción de la red. Es decir, las redes se clasifican por los componentes físicos que tienen. Así, se pueden definir los siguientes tipos de red:

- Coaxial : Son las redes que recurren a un cable constituido por un núcleo de cobre aislado por plástico y de un recubrimiento metálico; estos cables son similares al que se utiliza para conectar una televisión.
- Par trenzado : El cable que utilizan es parecido al que se emplea en la conexión de teléfonos.
- Fibra óptica : Consistente en cableado realizado con base en una combinación de vidrio y materiales plásticos de recubrimiento.
- Microondas : Las redes que utilizan microondas no utilizan cables para conectarse unas con otras. Su forma de comunicación se asemeja a las transmisiones de radio, por lo que se necesitan antenas para emitir o recibir una señal. Se utilizan básicamente para unir equipos a nivel metropolitano.
- Satélite : Estas redes utilizan la comunicación vía satélite. Al igual que las anteriores, no requieren de cables entre computadoras. Son utilizadas para unir equipos a nivel metropolitano o mundial.

Aunque esta clasificación es muy precisa, las redes varían según la tecnología que se emplea para unir los equipos remotos. Así, una red puede caer en más de una clasificación, ya que en su construcción es posible que haya más de un tipo de enlace.

### Por su Topología

Esta clasificación es útil por la representación gráfica que se proporciona de los componentes de una red. Para simplificar el diseño de los equipos, éstos se representan como puntos (nodos) y los medios de comunicación mediante líneas. Mediante la topología diferenciamos a las redes según el esquema de trabajo con el cual se construyen. Así, tenemos los siguientes tipos de redes:

## 1. Conceptos básicos

- Bus : Apareta una red de distribución lineal con varias tomas para los diferentes equipos que están conectados. Ver Figura 1.4A.
- Anillo : Se basa en el mismo principio que la red anterior, con la diferencia de que la estructura que forma es cerrada, dando la apariencia de un círculo. La información en estas redes viaja en un solo sentido, por tal motivo son susceptibles de fallar cuando un equipo lo hace. Ver Figura 1.4B.
- Doble anillo : Su funcionamiento es parecido al de la red anterior, variando en que la información viaja en ambos sentidos y en que cuenta con la capacidad de evitar aquellos equipos que bloquean el flujo de información. Ver Figura 1.4C.
- Estrella : Para este tipo de redes existe un nodo central por medio del cual se comunican. Su funcionamiento es muy parecido al de los teléfonos conmutados. Ver Figura 1.4D.

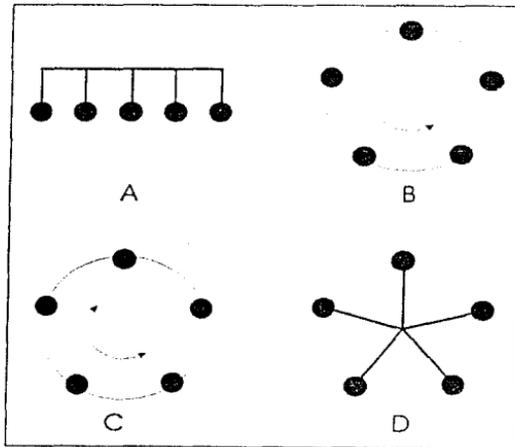


Figura 1. 4 Topología de redes

Como vemos, las diferentes clasificaciones nos permiten conocer más sobre aspectos particulares de cualquier red. La combinación de la extensión, los medios de conexión y la topología de cada red en particular, permite cierto manejo y disposición de información en ellas.

## 1. Conceptos básicos

### 1.3 Protocolos de comunicación

Se define como protocolo de comunicación al conjunto de reglas y estándares que se deben seguir para establecer una comunicación entre los distintos equipos que se conectan a una red.

Cada fabricante crea equipos y, de igual manera, especifica las reglas de comunicación (protocolos), que deben seguir los diferentes elementos que integran la red.

La organización americana ISO (*International Standards Organization*), definió una serie de reglas para establecer la comunicación entre dos o más equipos de cómputo para formar una red denominada *Modelo de Referencia OSI (Open System Interconnect)*, el cual consiste en una serie de capas que definen la comunicación entre computadoras. Cada capa realiza funciones específicas que permiten a las aplicaciones su ejecución en sistemas de arquitectura distinta, es decir, comunicarse unas con otras como si estuvieran operando en el mismo sistema<sup>1</sup>. Dichas capas se representan en la Figura 1.5:



Figura 1. 5 Modelo OSI

<sup>1</sup> ISO-OSI model (<http://studsys.mcs.mu.edu/~chong/mcs210/iso-osi.html>)

Cada una de las capas del Modelo OSI tiene una función determinada en el funcionamiento de una red. A continuación se presentan las funciones específicas de cada capa :

- **Capa 1 (Física) :** En este nivel se definen las señales eléctricas o electromagnéticas que se tomarán en cuenta entre dos puntos de la red. Se refiere al medio físico utilizado en la comunicación.
- **Capa 2 (Transmisión de datos) :** En este nivel se define la transmisión de datos entre dos puntos de la red. La información es dividida en bloques denominados "paquetes", y se verifica la transmisión de éstos.
- **Capa 3 (Red) :** En este nivel se define el mejor camino a seguir entre un punto inicial y uno final. Estos puntos no necesitan ser contiguos. A través de esta capa se establece el camino a seguir por los paquetes de información, pero no se verifica que la información llegue correctamente. Aquí también se definen las direcciones IP<sup>2</sup> para cada máquina.
- **Capa 4 (Transporte) :** En este nivel se realiza el re-ensamble de paquetes de información y se elimina información duplicada, en caso de haberla.
- **Capa 5 (Sesión) :** Establece la sincronización de los mensajes para realizar una labor particular, es decir, asigna el tiempo a cada usuario para hacer uso de los recursos de la red, logrando una sesión ordenada.
- **Capa 6 (Presentación) :** Permite la comunicación entre computadoras de diferentes arquitecturas realizando las transformaciones adecuadas a los formatos de información.
- **Capa 7 (Aplicación) :** Se definen utilerías básicas que permiten el uso de la red.

Cabe aclarar que en este modelo cada capa da soporte a la siguiente.

Este modelo fue desarrollado a fin de proveer un marco para el entendimiento de la manera en que la información es enviada de una computadora a otra<sup>3</sup>. Sin embargo, el modelo OSI es sólo de referencia, ya que existen muy pocas redes que lo tienen implementado, por lo que se usa generalmente para investigación. Por tal motivo, este modelo funciona simplemente para realizar comparaciones y poder precisar el funcionamiento entre redes.

Como se observa, las redes tienen reglas para establecer la comunicación entre los diferentes equipos que la conforman. A través de dichas reglas es posible la implantación de servicios con fines específicos, tales como la transferencia de archivos e información, búsqueda de información, o comunicación entre usuarios, por citar algunos ejemplos.

<sup>2</sup>Una dirección IP (Internet Protocol) es la dirección que identifica de manera única a una máquina. Esta dirección se conforma de cuatro series de dígitos cuyo rango va de 0 a 255. Ejemplo: 132.248.27.10

<sup>3</sup> AT&T Guide -- ISO OSI Reference Model (<http://www.att.com/interpan/overview/atn00417.html>)

## 1. Conceptos básicos

Un conjunto de reglas de comunicación ampliamente usadas se encuentran definidas en los protocolos TCP/IP.

### 1.4 Los protocolos TCP/IP

TCP/IP (*Transmission Control Protocol/Internet Protocol*), es un conjunto de protocolos de comunicación desarrollados para compartir recursos entre computadoras a través de una red.

Esta serie de protocolos establecen las reglas de conexión en una red. En otras palabras, para que un equipo se pueda conectar a *Internet* dicho equipo debe contar con software de TCP/IP. Originalmente estos protocolos fueron diseñados para el sistema operativo UNIX, sin embargo, actualmente el software basado en TCP/IP está disponible para la mayor parte de los sistemas operativos.

Al igual que el modelo OSI, TCP/IP establece una serie de capas que regulan los aspectos de comunicación en una red. En la Figura 1.6 se listan las 5 capas de TCP/IP :



Figura 1. 6 Capas de TCP/IP

El objetivo de cada capa es el siguiente :

- **Capa de Hardware :** Es la capa física de la red que conecta a las computadoras. Esta capa corresponde a la capa 1 del modelo OSI.
- **Capa de Interfaz de Red :** Es responsable de comunicaciones a bajo nivel en la red. Este nivel consiste en una serie de dispositivos que controlan la transmisión y recepción de datos a nivel físico de la red. Corresponde a las capas 2 y 3 del modelo OSI.

## 1. Conceptos básicos

- **Capa Internet :** Provee comunicación entre dos computadoras; no sabe de programas de aplicación ni del contenido de los datos. Únicamente de cómo comunicarse con otras computadoras en la red. Esta capa corresponde a la capa 4 del modelo OSI.
- **Capa de Transporte :** Provee comunicación entre aplicaciones de computadoras. Es responsable de proveer un intercambio seguro de información. Esta capa puede soportar múltiples programas de aplicación al mismo tiempo; corresponde a la capa 5 del modelo OSI.
- **Capa de Aplicación :** Es un programa que utiliza el modelo TCP/IP para comunicación. El programa tiene la opción de comunicar mediante un flujo de bytes de datos o por mensajes individuales. Esta capa corresponde a la capa 7 del modelo OSI.

Hay que notar que la capa 6 del modelo OSI no tiene una capa definida en TCP/IP.

La representación de información y el conjunto de acciones sobre la información son definidos explícitamente en TCP/IP, a diferencia de OSI, donde son dinámicos y pueden ser modificados. La Figura 1.7 presenta una comparación más precisa entre los modelos OSI y TCP/IP<sup>4</sup>.

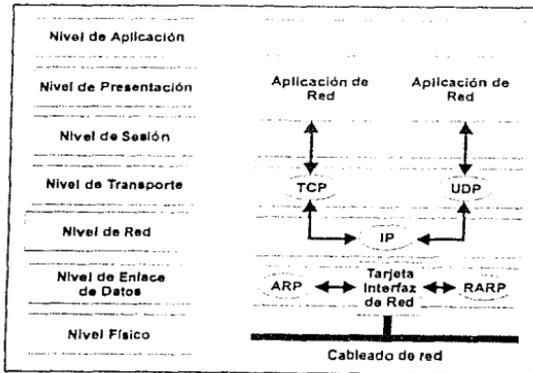


Figura 1.7 OSI - TCP/IP

Los protocolos TCP/IP tienen las siguientes características:

<sup>4</sup> Para la mejor comprensión de los términos que presenta la Figura 1.7 se anexa un glosario al final del trabajo.

## I. Conceptos básicos

- TCP : Garantiza una transmisión "*end-to-end*" de datos desde el emisor Internet a un receptor Internet. Secciona los flujos de datos grandes en pequeños "paquetes" (también conocidos como datagramas) y re-ensamblados cuando llegan al sitio receptor. Afortunadamente, este análisis y re-ensamblado son transparentes<sup>5</sup> para el usuario.
- TCP : Es un protocolo abierto, es decir, que el usuario de Internet no está obligado a contar con un *software* o *hardware* en particular para soportar el estándar TCP/IP.
- IP : Proporciona un esquema de direcciones para las máquinas en la red. Dicho esquema consiste en asignar una serie de números para identificar de manera única a una máquina.

Cabe mencionar que al conjunto de protocolos TCP/IP se les asigna la denominación "*connectionless*". Esto significa que la ruta desde el emisor hasta el receptor no está predeterminada. A lo largo del camino, los paquetes de datos pueden encontrar numerosos *routers*<sup>6</sup> que usan diversos métodos algorítmicos para determinar el siguiente punto intermedio mediante una dirección IP.

El número IP o dirección IP es una dirección empleada, principalmente, para distinguir de manera única a una computadora en *Internet*. Los números IP constan de 32 bits<sup>7</sup>, normalmente expresados por una serie de cuatro números decimales separados por puntos, representación conocida como **notación de punto decimal**. Cada uno de estos números representa un byte del número IP; por ejemplo: el número IP en notación de punto decimal 255.97.58.1, representa al número IP 11111111 01100001 00111010 00000001 en notación binaria. El protocolo IP garantiza que dicha dirección sea única, por lo que cada máquina que se conecte a *Internet* tendrá una dirección irrepetible dentro de ella.

### 1.5 La Internet

Hoy en día la *Internet* es conocida como "la red de redes". Esta acepción se debe al alto nivel de penetración que tiene en actividades tanto académicas como comerciales. Se calcula que en la actualidad la Internet da servicio a cerca de diez millones de personas en el mundo.

Parte de sus estándares actuales, así como su infraestructura se deben a la Agencia de Investigación de Proyectos Avanzados (ARPA - por sus siglas en

<sup>5</sup>La denominación "Transparente al usuario" quiere decir que este no se da cuenta de qué operaciones se realizan para obtener el resultado solicitado.

<sup>6</sup>*Router* es un dispositivo físico mediante el cual se pueden conectar dos redes de diferente topología.

<sup>7</sup>El bit es la unidad mínima de información.

inglés), del Departamento de Defensa de los E.U., la cual a mediados de los años 60's crea la red ARPANet.

El fin de esta red era la de mantener contacto entre distintos centros militares estratégicos de información, aún cuando alguno de ellos desapareciera al ser destruidos por el enemigo.

La arquitectura y los protocolos utilizados para la ARPANet tomaron su forma actual entre 1977 y 1979. La red *Internet* como tal se inició alrededor de 1980 cuando ARPA comenzó a convertir las máquinas conectadas a sus redes de investigación en máquinas con el nuevo protocolo TCP/IP. ARPANet rápidamente se convirtió en la columna vertebral del nuevo Internet y fue utilizada para realizar muchos de los primeros experimentos con el TCP/IP. La transición hacia la tecnología Internet se completó en enero de 1983, cuando la Oficina del Secretario de Defensa ordenó que todas las computadoras conectadas a redes de largo alcance utilizaran el TCP/IP. A su vez, la Agencia de Comunicación de la Defensa (DCA) divide a la ARPANet en dos redes separadas: una para fines de investigación (que conservó el nombre de ARPANET), y la otra para fines militares (MILNET), con lo que finalmente hace que la ARPANet se transforme en lo que hoy se conoce como *Internet*.

El éxito que tuvo la tecnología TCP/IP y la *Internet* entre algunos investigadores de ciencias de la computación, influyó para que la Fundación Nacional de Ciencias (NFS) propiciara su divulgación. En 1985 se comenzó un programa para establecer redes de acceso distribuidas alrededor de sus seis centros de investigación con supercomputadoras. En 1986, la NFS proporcionó fondos para la construcción de redes regionales, cada una de las cuales conecta en la actualidad importantes instituciones científicas de investigación en cierta área.

El crecimiento de *Internet* no se limitó a proyectos con fondos gubernamentales. Grandes corporaciones computacionales se conectaron a *Internet* así como otras de diversas áreas como compañías petroleras, automovilísticas, empresas electrónicas, compañías farmacéuticas y de telecomunicaciones.

### 1.6 Tipos de acceso a Internet

El acceso a *Internet* se puede realizar de dos formas :

- **Directa** : En este tipo de acceso, el equipo que se utiliza representa un nodo individual en la *Internet*, capaz de realizar cualquier cosa que sea posible hacer en la red. Dicha máquina tiene su propio número IP, y mediante ésta puede establecerse o utilizar un servicio de Internet. Este es el tipo de conexión más deseable.

## 1. Conceptos básicos

- **Indirecta** : En este tipo de acceso, el equipo final que se utiliza trabajará como una terminal<sup>8</sup> conectada a una computadora principal (algunas veces via módem<sup>9</sup>), la cual tiene acceso directo a la Internet. Por lo general, los servicios y dispositivos están restringidos por el dueño de la computadora principal. En general, esta es la forma más común de acceso

El uso de TCP/IP y los equipos que están conectados directamente a la Internet, han dado lugar a la creación de numerosos programas que permiten la integración de servicios que atienden necesidades específicas. En el siguiente capítulo se presentan algunos de los más utilizados en la red Internet.

---

<sup>8</sup> Se conoce como terminal a aquel equipo de cómputo que establece comunicación con otra máquina para hacer uso de algunos recursos de la misma.

<sup>9</sup> MODEM (MODulator, DEMolulator) – Dispositivo que conecta a una computadora con la línea telefónica, permitiendo a la primera comunicarse con otras computadoras a través del sistema telefónico.

# CAPÍTULO 2

## Servicios en *Internet*

### Servicios en Internet

En el capítulo anterior se analizaron los conceptos técnicos en que se basa la comunicación entre las computadoras que conforman la red *Internet*. A medida que el uso de ésta se fue expandiendo se tuvo la necesidad de crear servicios que atendieran las necesidades de sus usuarios.

En el presente capítulo se analizarán algunos de los servicios que presta *Internet*, desde los servicios básicos, hasta los denominados super-servicios<sup>10</sup>.

#### 2.1 Internet y el modelo cliente/servidor

El modelo cliente/servidor para un sistema de computadoras en red involucra tres componentes: el *cliente*, el *servidor* y la *red*. Un *cliente* es una aplicación que corre en la máquina del usuario final; el *servidor* es una aplicación que corre en la computadora que provee la información. La aplicación cliente puede ser personalizada para el sistema del usuario y actuar como interfaz<sup>11</sup> de ese sistema a la información provista por el servidor. El usuario puede iniciar una petición de información o una acción a través de un programa cliente. Esta petición viaja por la red hasta el servidor. El servidor interpreta la petición y toma una acción previamente programada. Esta acción puede incluir una búsqueda en una base de datos o cambiar la información contenida en un registro de la base. Los resultados de la transacción, si los hay, son enviados de regreso para ser desplegarlos por el cliente.

Toda comunicación de este tipo sigue un conjunto de reglas o protocolos, los cuales son definidos por el sistema cliente/servidor. En la Figura 2.1 se resumen estas relaciones, mostrando el flujo de una petición del cliente al servidor y la transmisión de información desde el servidor hasta el cliente. Cabe aclarar que un cliente puede acceder a varios servidores empleando los protocolos que tanto el servidor como el cliente entiendan.

<sup>10</sup> Los super-servicios son aquellos que permiten utilizar más de un servicio al mismo tiempo.

<sup>11</sup> Un conjunto de elementos ya sean físicos o lógicos que permiten la comunicación clara entre dos sistemas ajenos.

## 2. Servicios en Internet

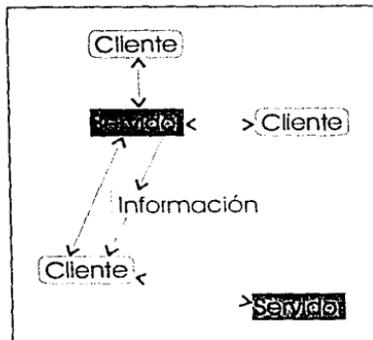


Figura 2.1 Arquitectura cliente/Servidor

La forma distribuida de las actividades de petición y respuesta del modelo cliente/servidor permite muchos beneficios. Dado que el cliente interactúa con el servidor de acuerdo a un protocolo pre-definido, el software cliente puede ser personalizado para la computadora del usuario en particular. El servidor no tiene que "preocuparse" por el hardware que el cliente utiliza, por ejemplo, un cliente de Web<sup>12</sup> puede ser desarrollado para computadoras Macintosh pero puede acceder a cualquier servidor Web. Este mismo servidor puede ser accedido por un "browser"<sup>13</sup> escrito para una estación de trabajo UNIX. Esto hace que el desarrollo de información sea más fácil, pues hay una clara demarcación de tareas entre el cliente y el servidor. Versiones separadas de la información no necesitan ser desarrolladas para una plataforma<sup>14</sup> en particular porque las personalizaciones necesarias están escritas en el software cliente de cada plataforma.

Una analogía para el modelo cliente/servidor es el sistema de emisión de televisión. Una persona puede comprar cualquier clase de televisor (cliente) para ver la emisión transmitida desde cualquier torre de televisión (servidor); ya sea que el usuario tenga un televisor portátil o una fija, ambas reciben información desde la estación emisora en un formato estándar y se despliega apropiadamente para cada tipo de televisor. Así, no son necesarios programas separados de TV para cada tipo de equipo, tal como para televisores a color, blanco y negro, o de diferente tamaño. Las estaciones de televisión que se crean tienen la capacidad de enviar señales a todos los televisores que se usen.

<sup>12</sup> Web es una forma abreviada de referirse al World Wide Web, mismo que será explicado más adelante.

<sup>13</sup> Browser se denomina a los programas clientes de Web.

<sup>14</sup> Plataforma es el conjunto de hardware y software que integran un sistema de cómputo.

Dentro de *Internet*, todos los servicios que en ella hay se basan en el modelo cliente/servidor. Siempre habrá un programa que el usuario manejará (cliente) y otro en espera de atenderlo (servidor) para cada servicio. En la mayoría de los casos, la parte del cliente y/o del servidor son transparentes para el usuario.

En las siguientes secciones se mostrarán algunos de los servicios más utilizados en *Internet*.

### 2.2 Telnet

Este servicio permite establecer sesiones de trabajo remoto en *Internet*. La sesión puede ser en una máquina cercana (local) o distante (remota). Cuando se establece una conexión remota, el usuario puede acceder a todos los servicios que esa máquina provee a sus terminales locales; puede realizar una sesión interactiva normal, esto es, teclear comandos o tener acceso a otros servicios, tales como búsqueda en catálogos de bibliotecas, acceso a periódicos y muchos otros.

En pocas palabras, Telnet permite el ingreso a una computadora conectada a red, aunque dicha máquina se encuentre en un sitio lejano. Una vez que el usuario se ha identificado ante el sistema, puede hacer uso de los recursos permitidos como si estuviese frente a ella.

La identificación de un usuario consta de dos partes : cuenta y contraseña. La cuenta es una clave que el administrador del servidor proporciona al usuario a fin de asignarle una "identidad" en un equipo. Dicha identidad sirve para asignarle recursos del equipo remoto: impresoras, unidades memoria, espacio en disco, etcétera. La contraseña funciona como llave de seguridad: sin ella no se puede acceder, aún conociendo el nombre de la cuenta, a ninguno de los recursos ni dispositivos que el servidor proporciona. La combinación de cuenta y contraseña debe ser única para cada usuario dentro de la máquina.

Siendo Telnet un servicio basado en el modelo cliente/servidor, cada parte tiene asignadas diferentes tareas, mismas que se describen a continuación :

El cliente :

- Crear una conexión TCP con un servidor.
- Aceptar lo que el usuario escriba en el teclado.
- Dar formato a la entrada de datos y transmitirla al servidor.
- Aceptar la respuesta del servidor.
- Dar formato a la salida para desplegarla de manera conveniente en la pantalla del usuario.

## 2. Servicios en Internet

El servidor :

- Informar al *software* de red que está listo para aceptar conexiones.
- Esperar una solicitud.
- Atender la solicitud.
- Enviar los resultados al cliente.
- Volver a entrar en el proceso de espera.

El programa cliente puede ser muy variado y depende en gran medida de la plataforma donde se desea utilizar (UNIX, Windows, Macintosh, Amiga, etcétera).

Un servidor debe ser capaz de manejar una variedad de clientes, ya que algunos correrán en el mismo tipo de computadora al servidor y otros no.

### Estableciendo una sesión

En los equipos con sistema operativo UNIX, el comando para establecer una sesión remota es **telnet**. La sintaxis de dicho comando es :

```
% telnet <dirección del equipo a conectarse>
```

Por ejemplo, si se desea establecer una sesión remota con un equipo cuya dirección sea 132.248.190.145 (o azteca.dgsca.unam.mx), la línea de comando será :

```
% telnet 132.248.190.145
```

o bien :

```
% telnet azteca.dgsca.unam.mx
```

A cualquiera de estas dos versiones se obtendrá lo siguiente :

```
Trying 132.248.190.145 ...  
Connected to 132.248.190.145  
Escape character is '^]'
```

```
UNIX(r) System V Release 4.0 (azteca)
```

```
!!!!!! Bienvenido usuario de azteca !!!!!!!
```

```
Le informamos que este servidor se encuentra en la Direccion General
```

## 2. Servicios en Internet

de Servicios de Computo Academico (DGSCA), en el Departamento de Evaluacion de Aplicaciones. La direccion Internet del servidor es 132.248.190.145 o azteca.dgscu.unam.mx.

Cualquier duda o comentario sobre este servidor favor de comunicarse a los siguientes telefonos (62)2-82-49 y (62)2-85-68, o las siguientes direcciones de e-mail :

alex@tztel.dcaa.unam.mx

Deseamos tenga una agradable sesion !!!

```
=====
login: alex
Password:
Last login: Tue Feb 11 11:41:27 from fobos.dgscu.unam
```

```

# # ##### ##### ##### ### ##
# # # # # # # # # ##
# # # # ##### # # # #
##### # # # # #####
# # # # # # # # # ##
# # ##### # ##### #### # # ##
```

Sun Microsystems Inc. SunOS 5.4 Generic July 1994

Azteca - BadBoy } } 1>

La primera parte corresponde al mensaje de bienvenida, en donde se pide al usuario se identifique con un *login* y un *password*. Si se proporcionan correctamente estos datos, el usuario puede hacer uso de los recursos que tenga permitidos en esa la máquina.

En equipos con interfaz Windows, existen programas tales como TELNET (de Microsoft), QVT/Term (de QPC software), que permiten establecer conexiones de tipo Telnet de forma más sencilla a través de menús y/o cajas de diálogo<sup>15</sup>.

<sup>15</sup> Una caja de diálogo es un espacio predefinido donde se pueden ingresar datos.

## 2. Servicios en Internet

Programa Telnet :

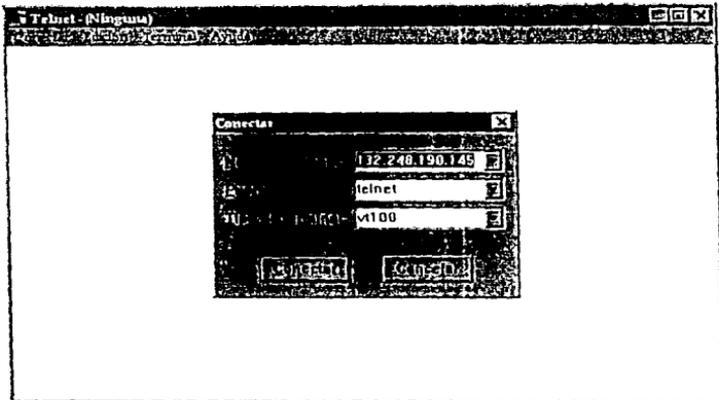


Figura 2.2 Programa TELNET

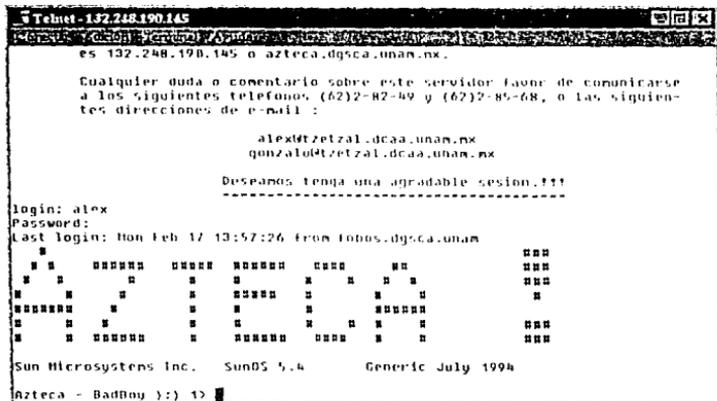


Figura 2.3 Programa Telnet

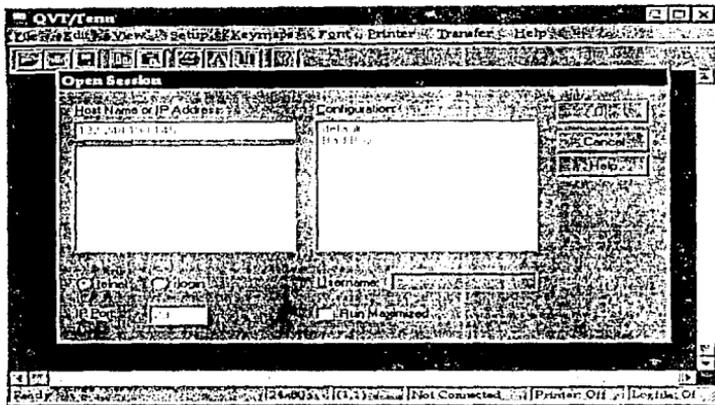


Figura 2.4 Programa QVT/Term



Figura 2.5 Programa QVT/Term



recibido algún mensaje. Si se desea escribir un mensaje a otra persona, se necesita conocer la dirección del buzón de esa persona.

Las direcciones de correo electrónico difieren de las direcciones de equipo en varios aspectos :

- El correo electrónico se envía a una persona, no sólo a una máquina.
- Algunas veces los nombres de personas se incluyen como comentarios en las direcciones de correo electrónico.

Si una persona desea tener una dirección de correo electrónico es necesario solicitarla con un proveedor de servicios de Internet<sup>17</sup>. Dicho proveedor le dará una clave y una contraseña en alguno de sus equipos. La contraseña sólo se utilizará para que el usuario pueda consultar de manera privada su correspondencia. Supongamos que una persona contrata una clave de correo electrónico con un proveedor; éste le da la clave "vperez" en una máquina cuya dirección es "buzones.correo.mx", la dirección de correo que obtuvo es la siguiente :

vperez@buzones.correo.mx

Como podemos ver, la dirección de correo se compone de dos elementos esenciales : la cuenta y la dirección del equipo donde se ubica. Es fácil ver que las direcciones de correo son únicas, ya que a los equipos se les asignan direcciones únicas. Si se diera el caso en que el nombre de una cuenta se repitiera en equipos diferentes, lo que obtendrá serán direcciones de correo electrónico diferentes.

Al igual que otros servicios, el correo electrónico cuenta con dos elementos claves para su funcionamiento : el servidor y el cliente.

El servidor realiza las siguientes funciones :

- Atiende la llegada de algún correo a las cuentas que tiene asignadas.
- Envía el correo a su respectivo destinatario.
- Envía mensajes generados por los usuarios de éste.
- Atiende las peticiones de los clientes.

El cliente, por su parte, tiene asignadas las siguientes funciones :

- Establece la conexión con el servidor de correo.
- Consulta por nuevo correo en el "buzón".
- Permite la consulta de los mensajes almacenados.
- Permite la creación de nuevos mensajes.

<sup>17</sup> La UNAM es un proveedor de servicios de Internet y cuenta con el servicio de correo electrónico.

## 2. Servicios en Internet

Como se ve, los servidores de correo electrónico son equipos que deben estar funcionando todos los días del año las 24 horas del día, ya que deben de permitir la llega de correo en cualquier momento. Por su parte, los clientes son usados para la consulta y emisión de correo.

Existe una gran variedad de programas cliente/servidor para el manejo de correo electrónico en Internet. La diferencia entre ellos radica en la forma en cómo realizan sus funciones y las opciones que como programa ofrecen. Actualmente la mayor parte de programas que se utiliza son los de dominio público (*freeware*), es decir, sin costo alguno.

Existen dos tipos de programas para enviar y recibir correo electrónico, los cuales dependen del *software* y tipo de acceso con que se cuente. Si se cuenta con un acceso a través del telnet, se está normalmente restringido a usar los programas existentes en la computadora a donde nos conectamos. Si se tiene acceso a Internet por medio de un equipo PC compatible o Macintosh, probablemente el manejo del correo se realice con un programa más atractivo, diseñado para ambientes graficos tales como Windows o System (de Microsoft y Macintosh, respectivamente).

Sin importar el tipo de programa que se utilice, el envío de mensajes se realiza prácticamente de la misma manera: se especifica el destinatario, se da un título al mensaje, se escribe el mensaje y se envía. Otras opciones que se pueden presentar son las de envío a varias personas, o la asociación de archivos para los destinatarios.

Toda esta información se tiene que proporcionar a los programas, los cuales presentan al menos los siguientes elementos para tal efecto :

- Línea "TO :": En esta línea se escribe la dirección del destinatario. Si se desea enviar a más personas, se necesita separar las direcciones de los destinatarios por una coma entre cada uno de ellas.
- Línea "CC :": CC es la abreviatura de "copy carbon", es decir, una copia del mensaje para una o más personas. Al incluir alguna dirección en la línea "CC :": se enviará el mensaje a las direcciones que se especificaron. La razón de esta opción es como la de los documentos oficiales que incluyen la línea "c.c.p." (con copia para), y se usa para dar niveles a las personas que reciben el mensaje.
- Línea "ATTACHMENTS :": Como en una carta o memorándum, se pueden anexas más mensajes u otros elementos. En el correo electrónico esto significa unir un archivo al mensaje, por lo que en esta línea se proporciona los nombre de los archivos que se van a enviar junto con el mensaje. Una vez que el destinatario recibe el mensaje,

puede desprender los archivos que estaban unidos al mensaje para utilizarlos según la información que contengan. Cabe mencionar que esta línea puede tener otro nombre dependiendo del programa que se emplee.

- Línea "SUBJETC ": En esta línea va un título para el mensaje y es propio de la persona que lo envía.
- Área del mensaje : Aquí va propiamente el mensaje. No hay reglas que regulen lo que se escribe y el usuario puede abordar los temas que desee.

Por ejemplo, supongamos que necesitamos enviar el siguiente mensaje a la dirección `vperéz@buzones correo mx`:

Hola Victor,  
  
Te recordamos que la reunión de la próxima semana será a las 13:00 hrs.  
  
Atte.  
Alex

A continuación se presentan varios programas con los campos llenos para enviar dicho mensaje:



Figura 2.7 Programa Pine

## 2. Servicios en Internet

Si los datos que se llenaron son correctos, el destinatario recibirá el mensaje que se le envió. Un aviso aparece en el programa del usuario (puede ser una ventana, una línea de texto, etcétera), indicándole la presencia del nuevo mensaje. La forma de organizar los mensajes es, por lo general, en "buzones". Existe un buzón en donde se almacenan los mensajes que llegan (en Eudora se llama IN y en Pine es INBOX), y otro para mensaje que se enviaron (OUT en Eudora y SENDBOX en Pine), además de los propios del programa de correo. Opcionalmente, pueden ser creados otros buzones para que el usuario organice su información.

Cualquiera que sea el programa cliente, el usuario sólo podrá ver el mensaje, mas no modificarlo. Cada mensaje que se reciba presenta la información según la persona que lo envió: dirección de correo electrónico, hora y fecha de envío, título del mensaje y el mensaje en sí.

A continuación se presenta un ejemplo conteniendo los datos arriba mencionados:



Figura 2.9 Programa Pine

### 2.4 Transferencia de Archivo - FTP (File Transport Protocol)

Para los equipos con sistema operativo UNIX, en un principio se estableció el protocolo FTP (*File Transport Protocol*) para la transferencia de archivos entre

## 2. Servicios en Internet

computadoras. Al igual que los servicios anteriores, funciona según el modelo cliente/servidor. En un inicio, era indispensable tener conocimiento de comandos básicos de UNIX para utilizar este programa cliente, FTP. En un sistema UNIX la sintaxis del comando FTP es el siguiente :

```
% ftp <dirección de la máquina>
```

donde la dirección de la máquina es la del equipo remoto.

En este tipo de cliente se utilizan comandos propios del servicio. A continuación se listan algunos de ellos :

Comando	Función
ascii	Definir una transferencia de tipo ASCII.
binary	Definir una transferencia de tipo binario.
dir	Listar los archivos y directorios.
get <archivo>	Efectuar la transferencia de un archivo de la máquina remota a la máquina local.
put <archivo>	Efectuar la transferencia de un archivo de la máquina local a la máquina remota.
mkdir <nombre>	Crea un directorio con un nombre específico en la máquina remota.
cd <nombre>	Cambia el nivel actual de directorio al que se especifica.
bye	Termina la sesión de FTP.

Actualmente existen versiones de clientes con interfaz gráfica, lo que hace más fácil la operación de transferencia de archivos.

Por lo general, un equipo servidor que provee el servicio de FTP designa una cuenta de usuario para cada una de las personas a las que deberá garantizar el acceso. A cada cliente (sea gráfico o no), se le pide proporcionar dicha cuenta y una contraseña cada vez que pueda utilizar el servicio de FTP. Las razones de estos requisitos de acceso son para proteger al equipo de usuarios desconocidos, ya que de otra forma, podrían modificar el contenido de las cuentas sin autorización de su verdadero dueño. Una vez que se tiene una cuenta con un servidor FTP, es posible copiar archivos hacia y desde el servidor cada vez que sea necesario.

Esta forma de control en los servidores FTP resulta poco conveniente para aquellos servidores destinados a distribuir *software* de dominio público, o bien, a compañías que ofrecen programas de prueba. Dichos sitios deben tener disponible el acceso para cualquier persona que desee conectarse a dichas máquinas mediante un cliente.

## 2. Servicios en *Internet*

La solución que se utiliza es crear un servidor FTP anónimo. Un servidor FTP anónimo tiene declarada una cuenta con el nombre de "*anonymous*" y la contraseña que solicita es la palabra "*guest*" o bien la dirección de correo electrónico del usuario. La cuenta anónima tiene restricciones de acceso para archivos y directorios de la máquina servidora, lo que resulta lógico ya que cualquier persona puede acceder a ella.

### 2.5 *Búsqueda y recuperación de información*

Como se puede apreciar, *Internet* contiene mucha información, la cual puede ser consultada por sus usuarios. Durante el desarrollo que la *Internet* ha tenido, han surgido diversos servicios de búsqueda de información; de entre los más poderosos y utilizados encontramos los siguientes :

- Gopher
- Archie
- WAIS

A continuación se describe cada uno de ellos:

**Gopher** : es un servicio de búsqueda de información a través de menús de temas. En todo servidor Gopher, existe un menú principal, mismo que va presentando los diferentes sub-temas en los cuales se encuentra organizada la información. Cada sub-tema puede estar a su vez dividido en otros temas o bien, presentar la información del tema de consulta. Este servicio se sigue usando en la actualidad, pero se le han agregado herramientas que agilizan la búsqueda de información, conocidas como Veronica y Jughead. Dichas herramientas trabajan en la misma forma: solicitan al usuario una cadena de caracteres relacionadas con el tema que se busca, presentando posteriormente las ocurrencias que fueron encontradas. La diferencia entre ellas radica en que Veronica realiza la búsqueda en varios servidores a nivel mundial y Jughead sólo en el servidor donde se solicitó.

**Archie** : busca nombres de archivos y/o directorios en sitios FTP anónimos; dicha búsqueda se basa en una palabra proporcionada por el usuario. Para utilizar este servicio se necesita un programa cliente, el cual tiene declarada una lista de servidores FTP anónimos para buscar en ellos. El resultado de una búsqueda con Archie es una lista que presenta la dirección del servidor con la ruta completa del nombre de archivo o directorio que encontró; si se desea recuperar el o los archivos que se localizaron, obviamente se requerirá el uso del servicio de FTP, utilizando los datos que proporcionó Archie.

**WAIS (Wide Area Information Service)** : es un servicio de búsqueda de información en bases de datos. Los clientes WIAS dirigen las búsquedas a índices de bases de datos contenidos en el servidor. Actualmente el servicio de WAIS se realiza por medio de WWW (World Wide Web).

Estos servicios, como la mayoría de otros, fueron creados por y para investigadores, ya que ellos fueron los primeros en usar *Internet*. Por tal motivo los clientes resultaban poco atractivos y difíciles de usar por personas en general. Actualmente, existen muchos programas clientes que utilizan la interfaz gráfica, y han facilitado al público en general su aprovechamiento, sin embargo, el gran auge fue hasta que se creó un nuevo servicio, el WWW, mismo que se presenta a continuación.

### 2.6 WWW (World Wide Web)

El *World Wide Web* (WWW), se origina en 1989 en el Laboratorio Europeo de Física de Partículas (CERN) en Ginebra Suiza. Tim Berners-Lees, egresado de la Universidad de Oxford, llegó al CERN y vio la necesidad de los físicos para colaborar en tiempo real, no sólo en uno, sino en varios proyectos. Con su experiencia en procesamiento de texto y comunicaciones en tiempo real, buscó crear un nuevo tipo de sistema de información en el cual los usuarios pudieran colaborar e intercambiar información durante el desarrollo de un proyecto.

Berners-Lee utilizó la tecnología de hipertexto<sup>18</sup> para enlazar todos los documentos de una red que pudiera ser recorrida de muchas formas para buscar información. En cooperación con otros investigadores del CERN, definió una arquitectura basada en la red *Internet* usando especificaciones abiertas y gratuitas, así como programas de muestra tanto para clientes como para servidores. Este equipo de trabajo desarrolló un "*Browser*" o Navegador<sup>19</sup> en modo texto<sup>20</sup>, el cual es el denominador más bajo entre los navegadores y puede ser usado desde casi cualquier tipo de terminal. Linux, un Navegador con interfaz "*full-screen*", fue desarrollado en la Universidad de Kansas. Aunque ambos Navegadores soportan el ambiente de hipertexto, no soportan elementos gráficos o multimedia.

El uso generalizado del *Web* no fue sino hasta 1993 con la versión del primer Navegador gráfico: *Mosaic*. Marc Andreessen, un estudiante de la Universidad de Illinois en Urbana-Champaign (UIUC), que estuvo trabajando en el Centro Nacional para Aplicaciones en Supercómputo en la construcción de herramientas de visualización científica, produjo *Mosaic*, un Navegador para el *Web* con una interfaz que permitía, mediante un clic a una liga, navegar en el *Web*, además de tener la capacidad para desplegar imágenes. A partir de entonces se añadió a *Mosaic* visualizadores externos que le agregan capacidades multimedia.

<sup>18</sup> Hipertexto es aquel texto mediante el cual se puede "saltar" dentro de la información.

<sup>19</sup> Aunque el término *Browser* no tiene una equivalencia propia en el idioma Español, se adopta el término "Navegador" como su traducción formal.

<sup>20</sup> Modo texto se refiere cuando un programa funciona sin utilizar algún ambiente gráfico (Windows, Solaris, etcetera), en otras palabras, sólo puede utilizar el juego de caracteres ASCII para desplegar información.

## 2. Servicios en Internet

En 1994 Marc Andeersen y Jim Clark forman la compañía *Mosaic Communications Corporation*, que posteriormente sería conocida como *Netscape*. Su navegador se ha extendido por el mundo de manera considerable, siendo al momento de desarrollar el presente trabajo el líder en el mercado.

Un competidor de *Netscape* que existe actualmente es *Microsoft Explorer*. Copia muchos de los estándares que ha impuesto *Netscape*, pero agrega muchas extensiones para los productos Microsoft: Word, Excel, Visual Basic, entre otros. Aunque su penetración ha sido constante, no ha podido desbancar de su puesto a *Netscape*.

Con el avance y expansión de esta tecnología, una nueva generación de Navegadores está apareciendo con las siguientes características :

- Reproducción de sonido.
- Reproducción de video.
- Visualizadores de realidad virtual (VRML).
- Ejecución de scripts.

Todo esto, manteniendo una fuerte vinculación entre los avances *hardware* y los programas cliente que se están generando.

Como ya se había mencionado, un servidor de WWW proporciona un medio más fácil y accesible a la información. Esta información se presenta al usuario a través de páginas generadas según las especificaciones del lenguaje para hipertexto HTML.

El HTML (*HyperText Markup Language*), es un lenguaje que define "marcas" en un documento para resaltar sus características y estructura, permitiendo su despliegue subsecuente en un Navegador.

El propósito del HTML es el de permitir que los documentos contengan vínculos electrónicos hacia otros documentos. Un documento puede crear un vínculo hacia otro archivo de texto, imagen, sonido o video mediante el protocolo correspondiente. Las reglas que se definen para hipertexto en la WWW son el protocolo de transferencia llamado HTTP (*HyperText Transport Protocol*)

Las acciones que se desarrollan por parte del cliente (Navegador) y del servidor para el servicio de WWW, son las siguientes :

1. El usuario proporciona al Navegador una dirección de alguna página de algún servidor WWW. O bien, hace clic en una palabra subrayada (hipertexto) para ir a otra página.

2. La solicitud de la página anterior es llevada a través de la *Internet* hacia el servidor WWW de referencia.
3. El servidor WWW localiza la página HTML solicitada utilizando el nombre de la ruta de acceso suministrado, y manda el archivo de regreso al cliente a través de la *Internet*.
4. El visualizador del Navegador almacena temporalmente el archivo en la máquina local, interpreta el contenido del archivo (con formato HTML) y lo despliega en la pantalla.

En resumen, el cliente solicita una página y el servidor la proporciona.

La Figura 2.10 esquematiza los puntos anteriores:

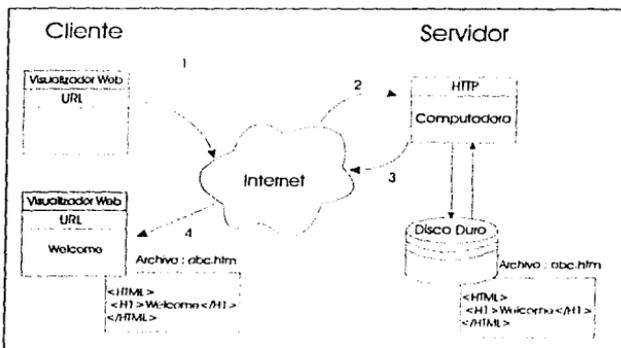


Figura 2.10 Arquitectura Cliente/Servidor para el servicio WWW

Las direcciones que se utilizan en el WWW se denominan URL (*Uniform Resource Locator*). Un URL es una camino estándar desarrollado para especificar la localización única de un recurso disponible electrónicamente<sup>21</sup>.

Los URL hacen posible que los clientes, u otros programas, puedan acceder a la información mediante los protocolos de los servicios disponibles en la Internet.

Un URL es como una dirección de correo: especifica la información necesaria para que alguien lo accese. Sin embargo, hay mucho más que eso, con los URLs se puede hacer referencia a una gran variedad de tipos de fuentes de información.

<sup>21</sup> A Guide URLs - <http://www.netSPACE.org/users/dwb/url-guide.html>.

## 2. Servicios en *Internet*

La sintaxis general de un URL es la siguiente :

<esquema>:<información-dependiente-del-esquema>

Un *esquema* indica el tipo de fuente de información que se está buscando y/o el mecanismo a usar para obtener ese recurso. Ejemplos de esquemas son "http", "gopher", "ftp" y "news"

La parte de *información-dependiente-del-esquema* incluye dos tipos de información : la dirección de la máquina en Internet que hace disponible a un archivo, y la "ruta" donde se encuentra ese archivo. Generalmente, el esquema se encuentra separado de la dirección de la máquina mediante dos diagonales (//), seguidas de la dirección de la máquina, y separada de la ruta completa del archivo mediante una sola diagonal (/).

Ejemplo :

<http://www.netscape.org/users/dwb/url-guide.html>

Donde el esquema "http" es para el Protocolo de Transferencia de HiperTexto. La dirección de la máquina es "www.netscape.org", y la ruta al archivo url-guide.html es "users/dwb/url-guide.html"

Asociadas a las direcciones se localizan las páginas WWW. Una página WWW es como una página de un libro o una revista; pero en vez de listas de información en formato de texto, párrafos y gráficas se despliegan todo el tiempo.

Para visualizar una página es necesario proporcionar un URL al Navegador; que además de tener la capacidad de desplegar páginas propias de Web con elementos multimedia en ella, se puede utilizar para acceder los diferentes servicios que ya existían en la *Internet* antes del servicio WWW. Para poder hacerlo, basta con dar el URL correspondiente para lograrlo. En el caso de Netscape, esto se logra con el menú "File" y la opción "Open Location" que presenta una ventana con una caja de texto en donde se deberá escribir el URL de la fuente requerida, tal como se ve en la Figura 2.11.

Usando el Navegador de Netscape y dando como referencia la dirección anterior, se podrá apreciar la página de la UNAM<sup>22</sup>, la sección "Netsite" indica el URL de la página actual (Figura 2.11).

---

<sup>22</sup> El contenido de la página así como su URL que se presentan pueden variar de la versión actual.



Figura 2.11 URL de la página actual

Existen más tipos de URLs para acceder a diferentes páginas o fuentes de información, cada uno con su propia sintaxis. A continuación se presentan la sintaxis y ejemplos de algunos de ellos.

#### HTTP (HyperText Transfer Protocol)

`http://<Servidor>[:puerto][<Ruta>[?<Parámetros>]]`

Para el HTTP, el *Servidor* es la dirección de servidor WWW, y el *puerto* es el número del puerto que se especificó para el funcionamiento de este servicio. En la mayoría de los casos, se puede omitir el puerto (siempre que se ponga deberá ir precedido por dos puntos), siendo 80 el puerto predeterminado para estos casos. La *Ruta* indica al servidor WWW cuál es el archivo que se busca; si se omite, indica que se busca la "página principal" para el sistema. Los *Parámetros* pueden ser usados para enviar información al servidor, generalmente a un programa ejecutable (CGI).

## 2. Servicios en Internet

### FTP (File Transfer Protocol)

**ftp:// < cuenta> : < contraseña> @ < Servidor> [: puerto] / < Ruta>**

El *Servidor* es la dirección de servidor FTP, la *cuenta* es la clave que tiene el usuario en ese servidor, la *contraseña* es la identificación del usuario y, en caso de ir estos dos datos se tiene que poner una arroba (@) para después indicar el servidor al cual acceder. Para FTP anónimos se puede omitir tanto la cuenta como la contraseña. El *puerto* es el número de puerto que se especificó para el funcionamiento del servidor FTP. En la mayoría de los casos, se puede omitir el puerto (siempre que se ponga deberá ir precedido por dos puntos). Para este servicio, el puerto por determinado es el 21. Finalmente, la *Ruta* indica al servidor FTP cuál es el archivo que se busca; si se omite, indica que se desea un nivel de directorio en ese servidor.

### Gopher

**gopher:// < Servidor> [: puerto] / < Ruta-gopher>**

El *Servidor* es la dirección de servidor Gopher, el *puerto* es el número del puerto que se especificó para el funcionamiento del servidor Gopher. En la mayoría de los casos, se puede omitir el puerto (siempre que se ponga deberá ir precedido por dos puntos). Este servicio establece al puerto 25 como el predeterminado para otorgar dicho servicio. La *Ruta-gopher* indica al servidor qué información se busca; si se omite, se está indicando que se desea el menú principal de ese servidor Gopher.

Hay que notar que cuando se usan URLs de FTP, HTTP, y Gopher la *Ruta-completa-del-archivo* puede algunas veces terminar con una simple diagonal. Esto significa que el URL en cuestión no especifica un archivo, pero sí un directorio. Un ejemplo de esto último es el URL de la página WWW de la UNAM, el cual es :

<http://www.unam.mx/>

En resumen, tenemos que para cada elemento y servicio en la Internet se tiene un URL asociado a él.

Un Navegador consta, en general, de las secciones identificadas en la Figura 2.12.

## 2. Servicios en Internet

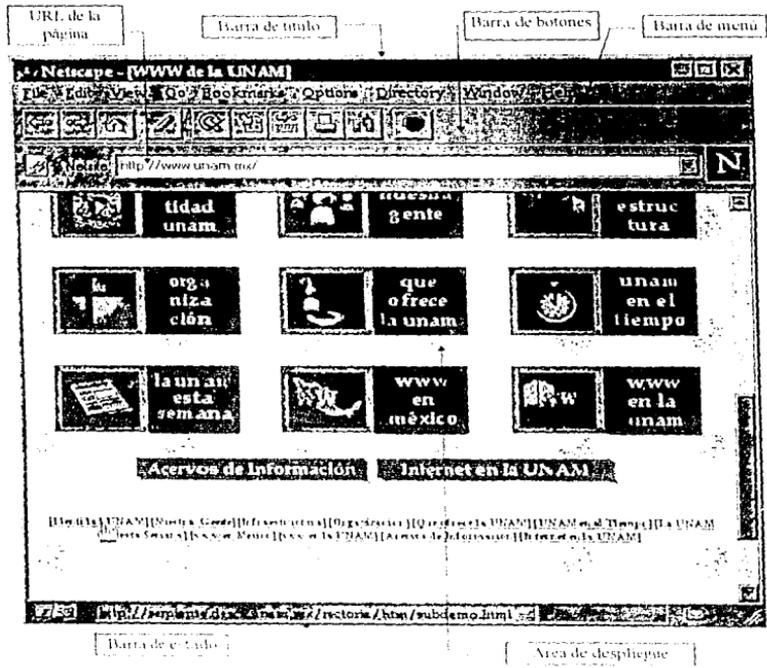


Figura 2. 12 Secciones de un browser

Los Navegadores actuales cuentan con una gran cantidad de elementos que permiten hacer una mejor presentación de la información permitiendo al mismo tiempo la creación de servicios muy variados a través de las páginas Web. Muchas instituciones (educativas, gubernamentales, iniciativa privada, etcétera), hoy en día ya cuentan con una presencia en el WWW. Cada día, un mayor número de personas tienen acceso al WWW y por consiguiente a muchos servicios, tanto existentes, ya no sólo los de búsqueda de información.

Las siguientes son reglas básicas sobre buenas páginas en el WWW :

## 2. Servicios en *Internet*

- Una página no debe tener un gran número de ligas. En vez de ayudar a la consulta de información la entorpece. En todo caso, será mejor crear un índice general.
- Una página no debe tener muchas imágenes en una página. Como cada imagen requiere ser transferida del servidor al cliente, la transferencia de información se hace más lenta. En todo caso, habrá que plantear un mejor diseño.
- Hay que tener cuidado con las diferencias culturales. Como el servicio de WWW es a nivel mundial, muchos contenidos pueden ser mal vistos por personas de otros lugares del mundo.
- Nunca hay que utilizar una liga que tenga de título : "Haga clic aquí". Esto resulta una redundancia, ya que cada liga de hipertexto lo indica.

Estas reglas definen un marco general de diseño de páginas individuales; cuando se crean varias de ellas enfocadas a un tema en particular, es decir un sitio, se deben considerar un mayor número de aspectos.

Para lograr un sitio profesional, su desarrollo involucra una gran participación de profesionales de diversas áreas, así como la definición de objetivos y características que se desea tenga en el futuro. En el siguiente capítulo se plantea una metodología de desarrollo de sitios WWW que involucra muchos de los aspectos antes mencionados.

# CAPÍTULO 3

Metodología para el  
desarrollo de sitios  
WWW

### 3. Metodología para el desarrollo de sitios WWW

Como se definió en el capítulo anterior, un sitio para el *Web* combina texto, gráficas e hipermedios. Por tal motivo, los desarrolladores pueden utilizar un gran conjunto de conocimientos que sobre la presentación de información se han venido dando desde tiempos remotos. En este capítulo se presenta un desarrollo metodológico para la creación de un sitio para el *Web*, el cual se basa en las características y cualidades deseadas del sitio, además de las necesidades, experiencias del usuario, y el proceso de comunicación. La clave en esta metodología es que se esfuerza continuamente por desarrollar y mejorar las estructuras de información al conocer las nuevas necesidades de los usuarios.

#### 3.1 Presentación de la metodología<sup>21</sup>

La presente metodología de desarrollo de sitios para el *Web*, contiene muchos de los elementos del desarrollo tradicional de sistemas de información, pero los procesos son más abiertos, ya que el producto final (un sitio operativo) no debe ser tan estático como los medios tradicionales de información.

Los elementos y los procesos en un sitio *Web* están interconectados, y las decisiones que los desarrolladores hagan dependen de dichas interconexiones. La Figura 3.1 presentan esquemáticamente dichos elementos y procesos.

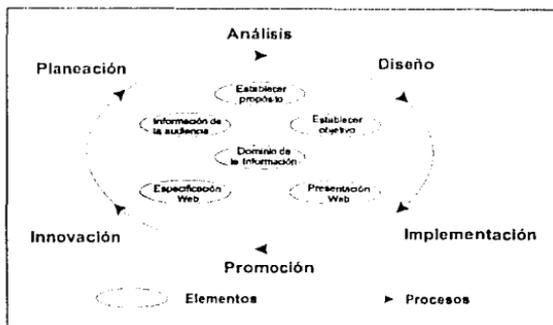


Figura 3.1 Metodología para el desarrollo de sitios

<sup>21</sup> El presente trabajo se basa principalmente en el realizado por John December y Mark Ginsburg en su obra titulada "HTML & CGI UNLEASHED" publicado por la editorial SAMS NET en 1995.

### 3. Metodología para el desarrollo de sitios WWW

La metodología tiene la característica de hacer que si algún elemento o proceso es débil, otro elemento o proceso más fuerte puede compensarlo. Por ejemplo, una buena implementación puede algunas veces mejorar un mal diseño. Un buen establecimiento de objetivos puede enriquecer un pobre propósito previamente establecido. Sin embargo, la meta no es tener esas debilidades, por lo que se busca es contrarrestar los inevitables problemas que se presenten.

#### 3.2 Elementos del desarrollo de sitios Web

Los seis elementos en la metodología de desarrollo de sitios para el Web son los siguientes :

- **Información de la audiencia** : es un conjunto de conocimientos sobre qué desea la audiencia del sitio, y cómo usa la información la audiencia actual. Esta información incluye la experiencia, los intereses, la propensión y todo detalle relevante de la audiencia, a fin de atender las necesidades del usuario. Esta información puede no estar completa en algún momento durante el proceso de desarrollo, por lo que deberá construirse un almacén de información en ese momento. La información que la audiencia requiere puede ser muy útil y precisa al mismo tiempo; con estos elementos, se puede diferenciar los distintos tipos de usuarios que accesan al sitio.
- **Establecer el propósito** : se basa en la razón de la existencia del sitio y de la amplitud del mismo. En todo momento un desarrollador debe tener un propósito establecido. Éste puede estar planteado en términos generales, o ser muy específico. Dicho propósito es dinámico en sí mismo; cada vez que una organización crea un sitio para establecer "presencia en el Ciberespacio"<sup>24</sup> puede querer hacer más específicos sus objetivos. Un establecimiento conciso de este propósito, sirve generalmente como una guía para el proceso de desarrollo.
- **Establecer el objetivo** : surge al establecer el propósito y define qué metas específicas debe realizar el sitio. El establecimiento del objetivo es dinámico, dado que puede volverse necesario definir otros más, así que el objetivo cambiará tanto como el propósito cambie, pero también lo podrá hacer en base a la información que sobre la audiencia varíe.
- **Dominio de la información** : es el conjunto de conocimientos e información que sobre un tema específico pretende abarcar el sitio, conformado tanto en términos de la información provista por los usuarios, como por aquella que los

<sup>24</sup> Ciberespacio .- Se usa para describir en su totalidad al rango de fuentes de información disponibles en las redes de computadoras.

### 3. Metodología para el desarrollo de sitios WWW

propios desarrolladores necesitan. Aunque no toda esta información puede estar disponible para los usuarios del sitio, éste dominio es, en la mayoría de los casos, esencial para los desarrolladores

- **La especificación del sitio** : es una descripción detallada de las limitantes y elementos que conformarán al sitio. Esta especificación indica qué partes de la información deberán ser presentadas al igual que cualquier limitante que se pudiera tener en la presentación de la misma. La especificación, como todos los otros elementos del sitio, puede estar en constante flujo y modificación.
- **La presentación Web** : es el medio por el cual la información es consultada por el usuario. La presentación es el resultado de los procesos de diseño e implementación que se desarrollan en base a la especificación del sitio. Durante dichos procesos, las selecciones se realizan a la par de las técnicas de diseño y de presentación con el fin de alcanzar las especificaciones del sitio

La lista de elementos involucrados en el desarrollo metodológico de un sitio *Web* muestran que hay muchas iteraciones y relaciones entre ellos, pues todos los elementos dependen de la mejor información que esté disponible sobre los demás para que el proceso de desarrollo sea exitoso. Por ejemplo, un desarrollador, al estar diseñando una parte del sitio, necesita saber si el objetivo es vender módems o capacitar a las personas sobre tales dispositivos. De manera similar, los elementos y procesos de la presente metodología de desarrollo interactúan entre sí. Estos procesos se describen a continuación

#### 3.3 Procesos del desarrollo

Los seis procesos contemplados por la metodología son los siguientes :

- **Planeación** : es el proceso de selección de oportunidades competitivas para la comunicación a fin de que las metas para las que el sitio fue creado puedan ser configuradas totalmente. Estas metas incluyen anticipar y decidir los objetivos de la audiencia así como los propósitos y objetivos de la información. La planeación también se hace para dominar la información a través de un proceso de definición y especificación, para determinar qué información deberá ser recolectada, cómo será recolectada y cómo será actualizada. Un "planeador" anticipa las habilidades planteadas en la especificación del sitio así como las habilidades necesarias para construir partes específicas del mismo. El planeador también anticipa otras fuentes necesarias para apoyar el desarrollo y puesta en operación del sitio.
- **Análisis** : es un proceso de acumulación y comparación de información sobre el sitio y su operación ordenada para mejorar, en general, la calidad del mismo. Una operación importante en este punto, es la de examinar la información

### 3. Metodología para el desarrollo de sitios WWW

acumulada sobre cada uno de los elementos y procesos involucrados en el desarrollo del sitio.

- **Diseño** : es el proceso por medio del cual un diseñador, trabajando con las especificaciones, toma decisiones sobre cómo un componente del sitio debe ser construido. Este proceso involucra el propósito, la audiencia, el objetivo, el dominio de la información, y lo que se pretende con el sitio. Un buen diseñador sabe cómo utilizar los elementos definidos en la especificación de la manera mas flexible, eficiente y elegante. Esto afecta fuertemente al resto de los procesos y elementos en el desarrollo del sitio, así que el proceso de diseño requiere un minucioso cuidado pues sienta las bases de la implementación.
- **Implementación** : en este proceso se construye el sitio usando HTML (*HiperText Markup Language*) como herramienta. El proceso de implementación es quizá similar al desarrollo de *software*, ya que involucra usar una sintaxis específica para codificar estructuras del Web en un lenguaje formal en archivos de computadoras. Al final lo que se debe tener es el sitio en sí
- **Promoción** : es el proceso de manejar todas las relaciones públicas de distribución de un sitio. Esto incluye dar a conocer la existencia del sitio para comunidades "en línea" a través de publicidad, así como de negocios establecidos u otra información relacionada con otros sitios. La promoción puede involucrar el uso de estrategias de mercado específicas o crear modelos de negocios.
- **Innovación** : es el proceso de hacer que los procesos del desarrollo continúen y se mejoren. Esto incluye el monitoreo de tecnologías para que innovaciones recientes puedan ser incorporadas al sitio, así como encontrar formas creativas o únicas para mejorar los elementos que ya se tienen o atraer más audiencia. Innovar es también mejorar continuamente el desempeño y calidad del sitio, excediendo las expectativas de los usuarios.

Aunque la metodología descrita aquí con respecto al desarrollo de sitios para el Web no sea la adecuada para en todas las situaciones, sirve como base para observar muchas distribuciones del desarrollo de sitios. Los procesos y elementos presentados por la metodología, pueden variar para un proyecto en particular, dichas variaciones deberán ser contempladas y aprovechadas por el equipo de trabajo que desarrolla un sitio en particular.

Cada uno de los procesos que conforman la metodología presentan una serie de particularidades que deben ser atendidas para su mejor integración. En las siguientes secciones se analizarán con más detalle cada una de ellas.

## 3. Metodología para el desarrollo de sitios WWW

### 3.3.1 Planeación

El planear la puesta en funcionamiento de un servicio en la Internet (ya sea Telnet, correo electrónico, transferencia de archivos, etc.), involucra la consideración de muchos aspectos. Uno de los más importantes son los siguientes :

- Equipo y servidor
- Contenidos generales

#### Equipo y servidor<sup>25</sup>

Existen dos alternativas básicas : obtener un servidor propio o rentar espacio en otro servidor.

Tomar la decisión de levantar y configurar un servidor *Web* propio está lejos de ser fácil. Cuando se toma en cuenta la cantidad de tiempo y dinero que se puede gastar en este tipo de proyecto, no existe lugar para errores.

Hay una enorme diferencia entre poner su propio sitio *Web* y poner una página en algún servidor. Esto último, se refiere a que cualquier persona o institución puede pagar una cantidad de dinero para levantar una página dentro de un sistema de acceso en *Internet*.

Si se tiene una compañía fuerte, lista para aprovechar el "ciber-mercado"<sup>26</sup>, se tiene una buena razón para levantar un servidor *Web*. Las razones que una compañía debiera considerar para tener su propio sitio *Web* son muchas: incluir imagen, mercado, proveedor de servicios, y posiblemente la venta directa.

Los institutos educacionales, el lugar de nacimiento de la *Internet*, también tienden a tener sus propios sitios *Web*. Obviamente, estas instituciones tienen razones de peso para tener sus propios servidores *Web* : dar a conocer investigaciones, trabajos y descubrimientos, y compartir información entre centros o instituciones.

Diversas organizaciones profesionales o de investigación también lanzan sus propios sitios *Web*, tales como el Área de Seguridad en Cómputo (<http://www.super.unam.mx/seguridad/>), el Instituto Nacional de Investigaciones Nucleares (<http://www.inin.mx/>) y el Instituto Mexicano del Seguro Social (<http://www.imss.gob.mx/>). Numerosos políticos, organizaciones de beneficencia, activistas, religiosos, y otras organizaciones también tienen sus propios sitios

<sup>25</sup> Tomado del artículo SITE CONSTRUCTION de Eric Richardson, Revista *Internet World*, Abril 1996.

<sup>26</sup> Ciber-mercado .- Se denomina así a la oferta de bienes y servicios que existe en la Internet.

### 3. Metodología para el desarrollo de sitios WWW

Existen muchas razones que se deben considerar para escoger entre administrar un servidor propio, o bien, pagar un servicio para albergar un sitio. A continuación se muestra un resumen de las ventajas y desventajas de cada opción:

VENTAJAS	DESVENTAJAS
<b>Construir un sitio propio</b>	
<ul style="list-style-type: none"><li>• Control completo sobre el contenido.</li><li>• Resulta "fácil" recoger información de visitantes vía formas</li><li>• Capacidad para seleccionar herramientas de seguridad</li><li>• Capacidad para rentar espacio en el servidor</li></ul>	<ul style="list-style-type: none"><li>• La configuración del servidor puede tomar meses.</li><li>• Se debe administrar el servidor.</li><li>• Requiere un gran y dedicado equipo de administración.</li></ul>

Tabla 3.1 Ventajas y desventajas de construir un sitio propio

VENTAJAS	DESVENTAJAS
<b>Pagar por un espacio</b>	
<ul style="list-style-type: none"><li>• No es muy caro</li><li>• No requiere sistema de mantenimiento</li><li>• Requerimientos mínimos de personal</li></ul>	<ul style="list-style-type: none"><li>• Dependencia del proveedor del sitio sobre seguridad y mantenimiento</li><li>• Depende del proveedor para cambios de contenidos</li><li>• Puede no tener capacidad para implementar formas y programas<sup>27</sup>.</li></ul>

Tabla 3.2 Ventajas y desventajas de pagar por un sitio

En general, la decisión dependerá de las necesidades y el presupuesto con que se cuente. Si el sitio es pequeño y la expectativa de tráfico<sup>28</sup> es ligera, lo mejor sería rentar el espacio. Si el sitio es grande y activo - esto es, si mucha gente estará agregando datos y visitando el sitio - sería ventajoso tener un servidor propio.

Como se observa, para cada alternativa existen ventajas y desventajas. Cada una de estas razones debe ser analizada por el equipo de personas encargadas a fin de atender las necesidades particulares del sitio que se desea crear.

Otro aspecto importante a considerar cuándo se decide tener un servidor *Web* es el equipo dónde será instalado. Dependiendo de la plataforma existen diferentes tipos de servidores, los cuales presentan características particulares en cuanto al

<sup>27</sup> Conocidos como CGI's que se discutirán más adelante.

<sup>28</sup> Tráfico - Cantidad de información que se transfiere desde un equipo en red.

### 3. Metodología para el desarrollo de sitios WWW

desempeño, instalación, extensibilidad, seguridad que se puede implantar y mantenimiento. Cada uno de estos aspectos se describen a continuación<sup>29</sup>.

#### Desempeño

Se entiende por desempeño las características con que cuenta un equipo respecto a su rendimiento. Esto es, la capacidad de atender las peticiones de los clientes (cuantos puede atender a la vez, la velocidad de respuesta, tipo de procesador, etcétera). El sistema operativo que se utiliza en cada equipo es también un factor importante para el rendimiento de un servidor; un sistema operativo multi-usuario ofrece una base de desarrollo mejor que los sistemas mono-usuarios.

#### Instalación

Es el proceso en el cual se establecen las características de funcionamiento del servidor. Va desde la adquisición del *software*, configuración (según requerimientos del sistema donde se monte, etcétera), hasta su puesta en funcionamiento.

#### Extensibilidad

En este punto se analizan las características de crecimiento de la plataforma. Es decir, si el equipo puede ser utilizado para más de un servicio, si pueden ser ampliados los recursos, tanto en *hardware* como en *software*, y la capacidad de utilizar herramientas nuevas, entre otros aspectos.

#### Seguridad

Es uno de los aspectos más importantes que se toman en cuenta para el ofrecimiento de servicios en la Internet. En los últimos años, la seguridad en las transacciones de información se ha vuelto vital, a tal grado, que se cuenta con organizaciones que buscan y publican errores o "huecos" de seguridad en los programas y equipos utilizados en la Internet. Estos huecos de seguridad pueden ser, por ejemplo, el paso de comandos de sistema a través de programas a fin de obtener información vital del sistema (números confidenciales, contraseñas, etcétera), o bien, obtener privilegios sobre recursos del equipo.

#### Mantenimiento

El mantenimiento tiene por objetivo que los servicios colocados en el servidor, se mantengan trabajando en condiciones óptimas y se vayan adaptando a las

<sup>29</sup> La plataformas que se consideraron para el análisis son las siguientes : Macintosh, UNIX, Windows 95, y Windows NT.

### 3. Metodología para el desarrollo de sitios WWW

necesidades que van surgiendo durante su uso. Dicho mantenimiento comprende actividades de monitoreo del sistema, detección y corrección de algún mal funcionamiento, actualización de datos, etcétera.

A continuación se presenta la tabla de resultados que diferentes plataformas obtuvieron<sup>30</sup>:

Sistema Operativo	% Desempeño	% Instalación	% Extensibilidad	% Seguridad	% Mantenimiento
Macintosh	**	*****	.	*****	**
UNIX	*****	.	*****	**	*****
Windows 95	****	****	****	***	***
Windows NT	*****	****	****	***	***

Tabla 3.3 Comparación de plataformas

Donde,

- Malo.
- \*\* Regular.
- \*\*\* Bien.
- \*\*\*\* Muy bien.
- \*\*\*\*\* Excelente.

Una vez evaluados los tipos de plataformas y las características particulares de trabajo de cada una, hay que evaluar los resultados obtenidos por los programas servidores que existen para poner un sitio Web.

Existen básicamente dos tipos de servidores: los de dominio público y los comerciales. Cabe mencionar que existe un amplio predominio de los servidores públicos sobre los comerciales. En la Tabla 3.4 se muestra un análisis sobre las diferentes marcas de servidores y su porcentaje de uso.

% Servidores	Tipo de servidor
54%	NCSA
17%	CERN
8%	NetScape
7%	Apache
5%	WebSTAR/MachHTTP
1%	Website
<1%	GN
<1%	WinHTTPD
<1%	Purveyor

Tabla 3.4 Tabla de servidores<sup>31</sup>

<sup>30</sup> Tomado del artículo SITE CONSTRUCTION de Eric Richardson publicado en la revista Internet World de Abril 1996.

<sup>31</sup> Datos de la revista Soluciones Avanzadas, Año 4 N° 33, 15 de mayo de 1996.

### 3. Metodología para el desarrollo de sitios WWW

Los servidores de dominio público son generalmente desarrollados por instituciones educativas o no comerciales. Están ampliamente documentados y en algunos casos se ofrece su código fuente para ser compilado por las personas que los adquieren. Es necesario revisar toda la documentación para poder instalar un servidor de este tipo. También se pueden encontrar programas y herramientas de dominio público para la administración del servidor. Son rápidos, ofrecen diversos niveles de acceso a las páginas del sistema a través de contraseñas y nombres de dominio. Existen grupos de trabajo alrededor de estos productos para determinar su eficiencia y cualidades, así como para detectar fallas o huecos de seguridad en su funcionamiento.

Por otra parte, los servidores comerciales no están documentados salvo las indicaciones de instalación y configuración; algunos cuentan con interfaces gráficas para tal efecto, además de proveer un esquema de seguridad como la disponibilidad de criptografía de llave pública y el estándar DES<sup>32</sup> de transacciones de información. Todo accesorio o herramienta extra que el proveedor pueda ofrecer para la administración tiene un costo adicional.

#### Contenidos generales

En este punto se debe obtener cuáles son los elementos básicos con los que se deben formar el sitio. No se requiere una explicación detallada sobre el funcionamiento entero del sitio, sino qué puntos debe cubrir. Debe dejar claro cuál es el objetivo del sitio, el propósito, dónde se obtendrá la información que irá en el sitio, entre otros.

#### 3.3.2 Análisis

La información que se presentará al usuario es muy variada y dependerá de los objetivos particulares de cada módulo de información.

Al plantear una página principal, por ejemplo, ésta debe proporcionar al usuario un panorama general del sitio. Una página principal (o *Homepage*), permite al visitante recorrer las diferentes secciones en las que se dividió el sitio que consulta. Debe ofrecer información general, dando una visión completa de que lo se ofrece.

La página principal de la UNAM, por citar un ejemplo, es un buen modelo<sup>33</sup> (ver Figura 3.2).

<sup>32</sup> DES - Data Encryption Standard, es un estándar americano para el envío de mensajes cifrados vía red. Para mayor información se recomienda al lector la página de la organización de seguridad RSA en <http://www.rsa.com>.

<sup>33</sup> <http://www.unam.mx/>

### 3. Metodología para el desarrollo de sitios WWW

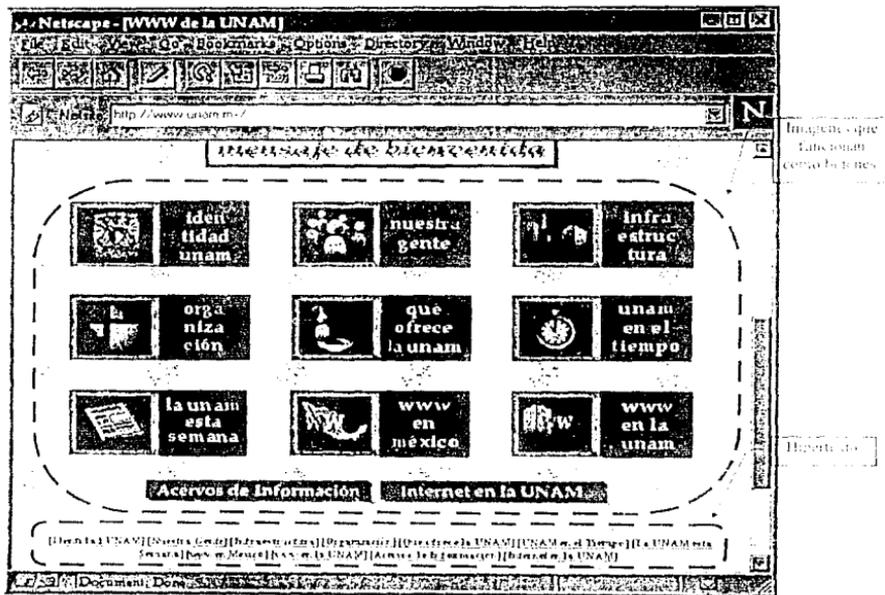


Figura 3.2 Página principal

Esta página presenta una sección para que el visitante conozca el sitio en sí, además de diversos aspectos de la UNAM, siendo una especie de menú.

Otros ejemplos se pueden tomar de sitios para buscar más sitios Web, por ejemplo, el caso del sitio Yahoo<sup>14</sup>, el cual presenta su página principal con una sección de temas, algo parecido al servicio de Gopher.

En este sitio, un aspecto importante es la pequeña sección en lo alto del índice de temas junto a un botón titulado "Search (Búsqueda)", que sirve para efectuar búsquedas de temas más específicos. La Figura 3.3 ilustra un sitio para búsquedas.

<sup>14</sup> <http://www.yahoo.com/>



Figura 3.3 Opción de búsqueda de Yahoo

Cuando el visitante proporciona una o más palabras clave, se iniciará la búsqueda. Este estilo de páginas es útil cuando se pretende ofrecer búsquedas dinámicas a través de una página Web. En el ejemplo, un visitante puede escribir "jornada" y al activar el botón "Search" presentará una página de resultados, según lo que haya encontrado en dicho sitio relativo al tema solicitado (ver Figura 3.4). Por regla general, la página de resultados aparece con hipertexto ligado a los documentos buscados.

### 3. Metodología para el desarrollo de sitios WWW

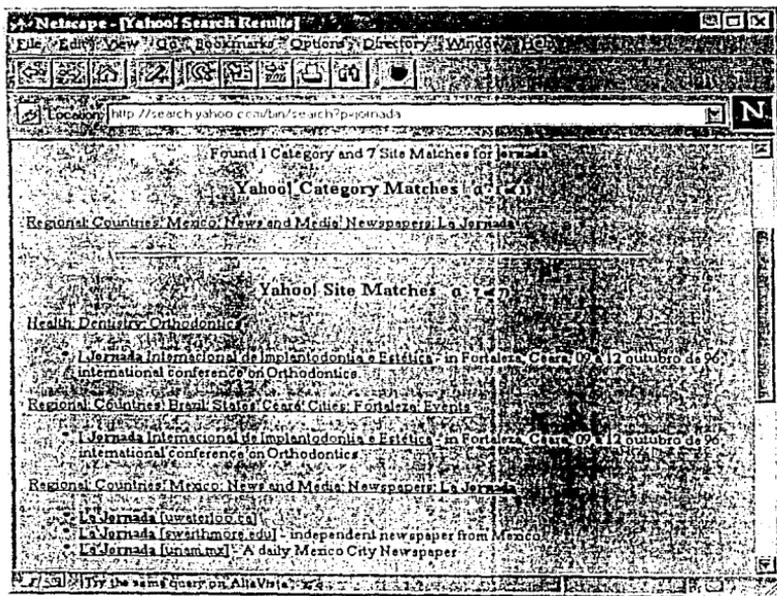


Figura 3.4 Resultado de una búsqueda con Yahoo!

Para sitios de distribución de software de prueba (*shareware*) o gratuito (*freeware*), las opciones de búsqueda son requeridas. Un ejemplo de este tipo de sitio es Tucows<sup>33</sup>. La página principal de este sitio es un directorio de otros servidores en el mundo, pertenecientes a esta organización, por lo que todos sus servidores tienen la misma información y archivos (ver Figura 3.5). La razón por la cual se distribuyan a nivel mundial, es que miles de usuarios los visitan para adquirir software nuevo. Si sólo se tuviera un sitio, traería como consecuencia la saturación de las líneas de comunicación. En este ejemplo en particular, sin un correcto análisis de necesidades de los usuarios, el sitio simplemente se perdería, ya que resultaría prácticamente imposible consultarlo.

<sup>33</sup> <http://www.tucows.com/>

### 3. Metodología para el desarrollo de sitios WWW

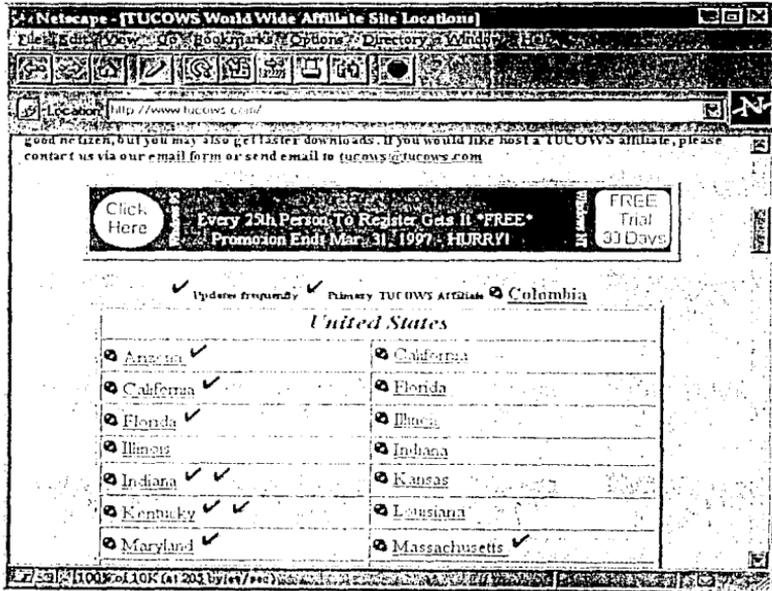


Figura 3.5 Página principal de Tucows

En la página principal de cualquiera de los sitios, se aprecia un nivel general de información el cual, mediante hipertexto, solicita al visitante seleccionar la plataforma para la cual desea obtener las copias del *software*: Windows 95, Windows 3.x, o bien Macintosh, además de una sección de búsqueda por palabra clave.

Cualquiera que sea la plataforma, en el sitio se observa una división de *software* para la búsqueda de programas (ver Figura 3.6).

### 3. Metodología para el desarrollo de sitios WWW

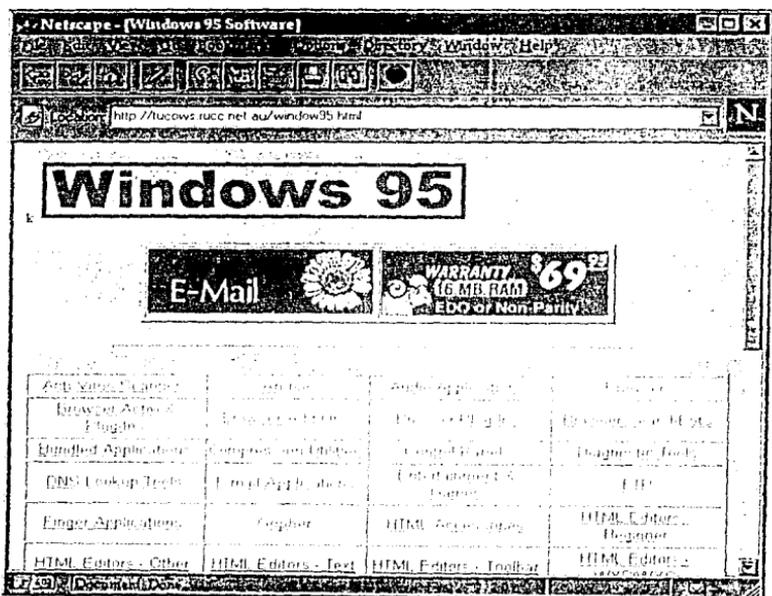
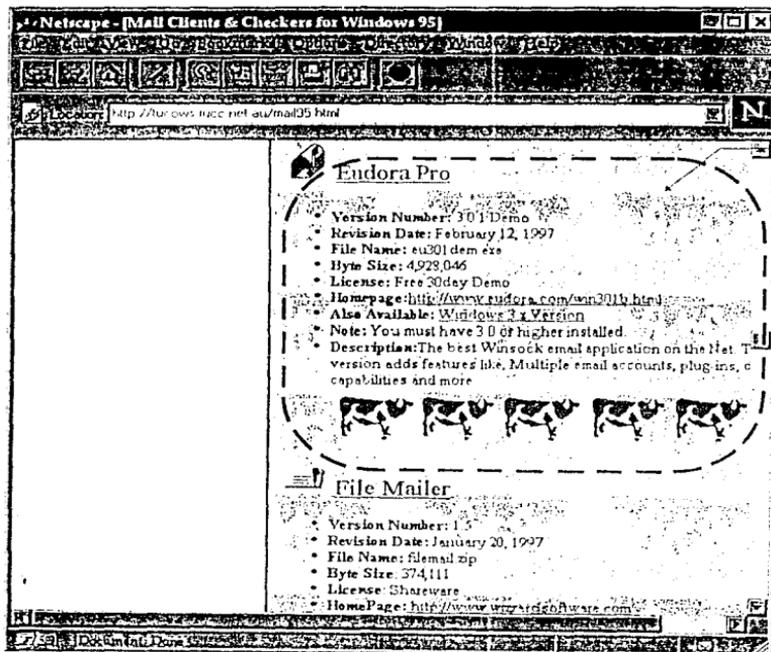


Figura 3.6 Clasificación del software

Si se selecciona, por ejemplo, la sección de E-mail Applications (programas para el manejo de E-mail), se presentará la lista de productos que permiten el uso del correo electrónico. Esta lista fue diseñada para que además de proporcionar la copia del software, el usuario sepa el tamaño del archivo, número de versión, una "calificación" al producto, etcétera (ver Figura 3.7). Todos los programas son presentados de la misma forma, con lo cual se proporciona una regla de navegación intuitiva y de fácil manejo.

### 3. Metodología para el desarrollo de sitios WWW



Información  
del programa

Figura 3. " Información del programa Endora Pro

Como se puede apreciar, gran parte del análisis de los elementos que intervienen en el desarrollo plantea una problemática a resolver por el equipo de trabajo.

Por lo tanto el análisis es un factor importante, ya que sirve de base para definir los elementos que intervienen así como de las soluciones que se pueden implementar en la creación de sitios Web.

El siguiente proceso plantea un diseño en base al análisis y conocimiento de los requerimientos y expectativas de los usuarios, que debe satisfacerlas apegiándose a las condiciones ya definidas.

### 3. Metodología para el desarrollo de sitios WWW

#### 3.3.3 Diseño

Un diseño de *Web* es esencialmente su vista y su funcionalidad. Un buen diseño deberá tomar en cuenta todos los elementos del análisis - información de la audiencia, definición del objetivo y propósitos, dominio de la información y la especificación del sitio - y combinarlos para producir un plan de implementación. Es así que los implementadores usan el diseño y las especificaciones del sitio para crearlo.

El diseñador recurre a técnicas de estructuración de la información, utilizando una o más metodologías de diseño lo que se conoce como paradigma computacional. Durante este proceso, el diseñador debe ser sensitivo con respecto a la experiencia de los usuarios en el uso de la información que está en el *Web* : texto, espacio y marcas<sup>36</sup>. Hay muchos resultados prácticos envueltos en el diseño, tales como consideraciones sobre imágenes y gráficas, cuánto cuesta realizar una página y cuáles imágenes o textos pueden ser hipertexto y cuales no.

Como no existe una sola forma de crear sitios, el diseñador puede escoger entre una gran variedad de enfoques mientras desarrolla uno.

Si un diseñador tiene una buena idea sobre qué contenido deberá tener el sitio, un método *Top-down* o *descendente* de diseño podría ser el mejor enfoque. En la metodología *descendente*, el diseñador comienza con una página principal para el sitio (frecuentemente llamada *Homepage*), y ramifica desde ahí. Se debe tener en mente crear páginas prototipo que contengan el mínimo de información pero que puedan servir al desarrollo del sitio. El beneficio del enfoque *descendente* es que el diseñador puede desarrollar páginas de acuerdo a la idea o propósito central. Esto provee una buena oportunidad para plantear la vista y funcionalidad del sitio, ya que todas las páginas se diseñan en base a la página principal. Una buena forma de hacer esto es mediante el diseño de un juego de plantillas de páginas, mismas que serán utilizadas durante el proceso de implementación.

Por otro lado, si el diseñador no tiene una buena idea de cómo será el sitio final (o simplemente cómo lo desarrollará), pero sabe cómo se especificaron tanto la funcionalidad como la presentación de las páginas, puede trabajar partiendo de las páginas secundarias hacia la página principal. Esto es particularmente cierto si se cuentan con algunas páginas ya elaboradas.

Si no se tiene alguna página con la cual comenzar, el diseñador puede elaborar páginas que cumplan con los objetivos especificados, para unirlos a través de páginas intermedias o hipertexto a la página principal. El beneficio es que el diseñador no se ve limitado por el estilo de la página principal, ya que puede diseñar páginas con estilo propio basadas en la función específica a cumplir.

<sup>36</sup> Marcas de hipertexto conocidas también como ligas.

### 3. Metodología para el desarrollo de sitios WWW

Posteriormente, el diseñador ajusta las páginas para crear una vista y funcionamiento general del sitio. Este enfoque se conoce como *Bottom-Up* o ascendente.

Un tercer enfoque que se utiliza es el de *Incremental/In-Time*. Con este enfoque se desarrollan páginas "al instante" cuando se requiere. Esto generalmente se hace cuando a partir de una página principal, se desea crear ligas hacia otras páginas que implementen objetivos particulares. Estas últimas son creadas y unidas con la premisa de que posteriormente se pueden agregar más páginas. Dicho enfoque es adecuado si los desarrolladores quieren tener un sitio de crecimiento rápido más que desarrollarlo por completo.

Una tarea específica en este diseño de sitios, es empaquetar<sup>37</sup> o presentar la información mediante bloques de modo que no abruman al usuario. Dicha información deberá ser listada en detalle en las especificaciones del sitio y verse reflejada en el objetivo establecido.

Existen muchas técnicas que auxilian en el proceso de separación de información. Como primer paso, para cualquiera de ellas, el diseñador debe recopilar aquellos documentos que contengan la información que será presentada en el sitio. Una vez obtenidas las fuentes de información, ésta debe ser dividida mediante los criterios establecidos en las especificaciones del sitio.

El resultado de dividir la información son los denominados diagramas de agrupamiento. Estos diagramas muestran las relaciones entre las piezas de información que el sitio deberá tener. Para obtener un diagrama de este tipo, se deberá comenzar por definir los elementos que existen en las perspectivas del objetivo, esto en términos de lo que el usuario espera y sin saturarlo de elementos de información. Como guía general, el número de piezas de información para atraer la atención de los usuarios en cualquier instante es de siete, más menos dos.

Supongamos que el diseñador cuenta con la siguiente información :

**Sitio** : Centro de Estudios Sobre Comunicación a través de Computadoras

**Propósito** : Ayudar a las personas a obtener información.

**Objetivo** : Proveer una lista de fuentes de información mediante hipertexto.

Un diagrama de agrupamiento puede servir como mapa para desglosar un sitio en paquetes. El diagrama correspondiente al problema planteado puede ser el que se muestra en la Figura 3.8.

---

<sup>37</sup> Un *paquete* es una página Web o grupo de ellas que están relacionadas.

### 3. Metodología para el desarrollo de sitios WWW

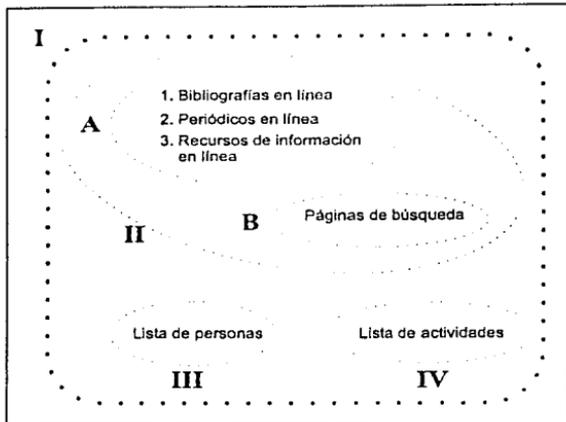


Figura 3. 8 Diagrama de agrupamiento

Con este desglose, el diseñador puede apreciar las relaciones que deberán tener las páginas dentro un paquete de información. Una vez identificadas, pueden ser creadas en el proceso de implementación que se verá posteriormente.

Eventualmente, cada paquete debe ser definido como una página de Web o como un conjunto de ellas.

De la figura del ejemplo anterior se pueden inferir fácilmente los siguientes paquetes de información :

- I. Un paquete conformado por otros tres paquetes : información sobre las personas, actividades y fuentes de información.
- II. Un paquete de recursos en línea, el cual contiene :
  - A. Un paquete de recursos de información en línea, que contiene :
    1. Un paquete de bibliografías en línea.
    2. Un paquete de periódicos en línea.
    3. Un paquete de recursos en línea.
  - B. Un paquete de búsqueda de páginas.
- III. Un paquete que contiene información sobre las personas.
- IV. Un paquete que contiene información sobre las actividades.

En este caso se definieron nueve paquetes.

### 3. Metodología para el desarrollo de sitios WWW

Una vez establecidos estos paquetes, se puede comenzar con la creación de las páginas que los forman.

Cuando el diseñador tiene un conjunto de páginas, el siguiente paso es el de especificar cómo se unirán todas ellas. El diagrama de agrupamiento es un buen comienzo hacia el análisis de ligas que deberán generarse. A partir de dicho diagrama se puede obtener un diagrama de ligas<sup>39</sup> inicial. Para obtenerlo se pueden seguir los siguientes métodos :

1. Unir las páginas mediante un englobamiento de paquetes derivada del diagrama de agrupamiento. Esta unión de páginas del mismo paquete se realiza mediante ligas. La Figura 3.9 muestra la utilización de éste método.

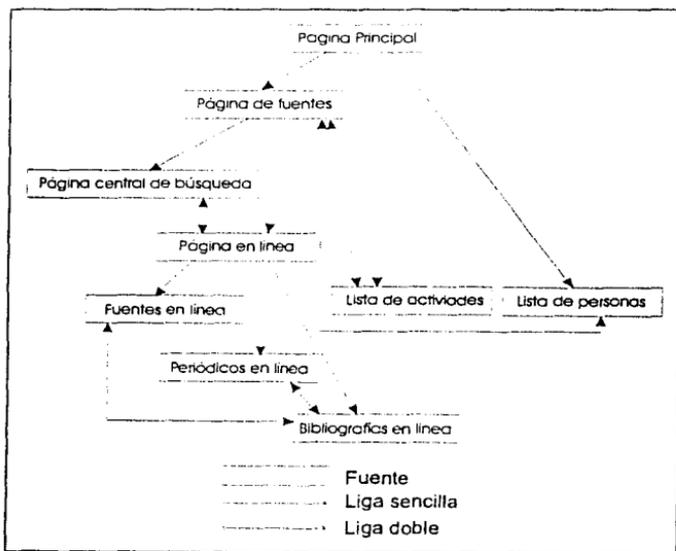


Figura 3. 9 Diagrama de Ligas (1)

<sup>39</sup> Este tipo de diagramas representan las relaciones que existen entre las páginas de un paquete, o bien, entre paquetes.

### 3. Metodología para el desarrollo de sitios WWW

2. Crear únicamente las páginas de los paquetes y unir cada una con las demás. Esto crea un sitio sin páginas subordinadas, en el cual todas las páginas de información están disponibles para las otras. La Figura 3.10 muestra la utilización de este método.

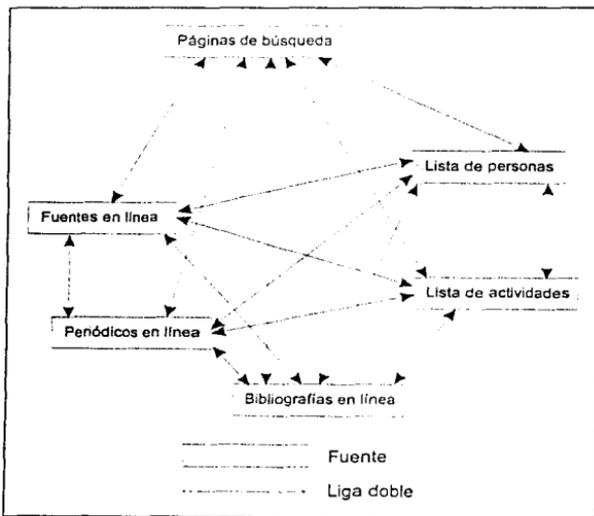


Figura 3.10 Diagrama de Ligas (2)

Otros métodos de unión de páginas que se pueden considerar son los siguientes :

- **Por necesidad** : Cuando por políticas establecidas, o bien por satisfacer las necesidades de un grupo particular de usuarios, se determinan que uniones se necesitan crear.
- **Por asociación** : Cuando existe una demanda particular por parte de los usuarios en una sección del sitio, se pueden establecer las ligas que permitan el mejor uso del mismo.

Existen un gran número de metodologías que ayudan al diseñador a definir la estructura del sitio<sup>39</sup>.

<sup>39</sup> Para consultar algunas técnicas de diseño se recomienda al lector la obra titulada "Teach Yourself HTML 3 in a Week" de Lemay, Laura editado por SAMS, 1996

### 3. Metodología para el desarrollo de sitios WWW

Aunque los paquetes, las páginas y los diagramas de ligas sean excelentes herramientas, el diseñador puede hacer uso de muchas otras técnicas para ayudar a la funcionalidad e imagen del sitio. Uno de estas es una "guía universal"<sup>40</sup> para el sitio entero, la cual es un diagrama que define la distribución y arreglo para el texto, la presentación y las ligas de una página en particular.

Por ejemplo, una guía universal se muestra en la Figura 3.11. El propósito de dicha guía es crear una plantilla para darle a todas la páginas del sitio una vista uniforme, lo cual ayuda a los usuarios del sitio a conocer qué elementos encontrará en cada página.

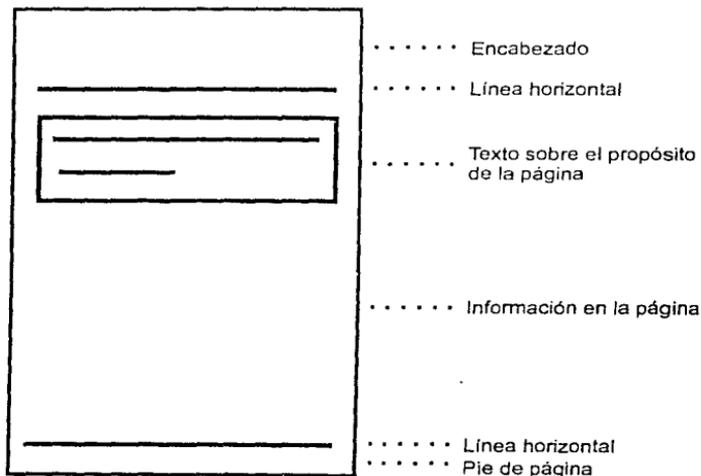


Figura 3. 11 Guía Universal (1)

Otra técnica para crear una imagen y funcionalidad unificados para el sitio es la de usar iconos<sup>41</sup> distintivos para representar tipos de información o un icono representando al sitio mismo. Estos iconos deben ser especificados en la guía universal. La Figura 3.12 muestra un ejemplo de este tipo, de técnica.

<sup>40</sup> Un diseño general que se debe respetar para todas las páginas de un paquete o sitio.

<sup>41</sup> Icono .- representación gráfica de alguna opción o característica propia de un diagrama.

### 3. Metodología para el desarrollo de sitios WWW

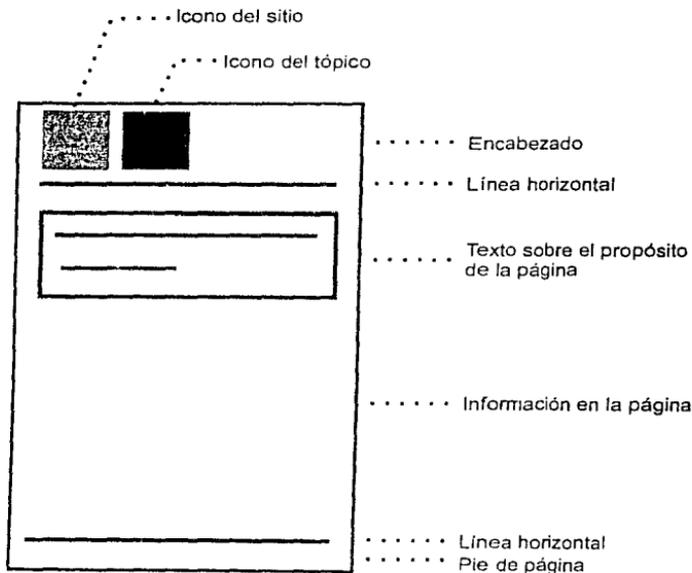


Figura 3. 12 Plantilla Universal (2)

Estas técnicas permiten al diseñador decidir cuál es la mejor opción que se apege a las expectativas del sitio. Una vez que se tiene un diseño propio, el siguiente paso es el de implementarlo. Como se mencionó, el proceso de implementación depende en gran medida del trabajo que el diseñador ofrezca y pueda enriquecer con su experiencia en el desarrollo de sitios para el Web.

#### 3.3.4 Implementación

La meta de la implementación es combinar los resultados de los procesos de planeación y diseño así como cualquier actualización o mantenimiento especificado como resultado del proceso de análisis para crear un sitio funcional. El trabajo del implementador o implementadores consiste en crear un puente entre esos procesos "abstractos" y las necesidades específicas de la implementación. Esta separación de procesos es útil a los implementadores ya que son libres de

### 3. Metodología para el desarrollo de sitios WWW

tomar las decisiones que necesiten hacer a su nivel de responsabilidad. Dicha separación también ayuda a los diseñadores y planeadores a enfocar su atención a las necesidades de la audiencia.

Un implementador debe recopilar a siguiente información, misma que ha sido generada por los procesos de la metodología :

**Productos del diseño :** Como resultado del proceso de diseño, fueron creados una vista, la funcionalidad, un juego de paquetes de información, diversas páginas y diagramas de ligas. Estos dos últimos productos son la mejor guía que el implementador usa en un trabajo.

**Especificación del sitio :** Como resultado del proceso de análisis, un analista especifica las características del sitio, sus limitantes u otros parámetros a seguir. Esta especificación guía al implementador en la toma de decisiones sobre muchas de las características del sitio.

**Especificaciones tecnológicas :** Aunque HTML es un estándar para la presentación de documentos en cualquier Navegador, muchos de ellos han implementado características propias. La competencia entre Navegadores ha creado como resultado una constante innovación en la cual los implementadores de sitios necesitan estar en constante actualización sobre el estado

**Requerimientos de actualización y mantenimiento :** Como resultado del proceso de análisis, se determinó el conjunto de elementos en HTML o actualizaciones en la información. Las actualizaciones pueden ser el resultado del proceso de planeación o del proceso de innovación mismo, que será explicado más adelante.  
del lenguaje HTML.

**Información proporcionada por los usuarios :** El implementador a menudo mantiene contacto estrecho con los usuarios a través del correo electrónico. Los usuarios pueden informar sobre ligas no actualizadas o proporcionar información específica sobre cómo trabaja el sitio. Algunas de estas recomendaciones serán irrelevantes, otras requerirán un poco de trabajo y algunas más necesitarán la revisión de los procesos de planeación, análisis o diseño.

El objetivo de un implementador será siempre el de crear y mantener el mejor trabajo posible. Para tal efecto sus principales tareas son las siguientes :

**Integrar** el diseño y cualquier otra información manejada previamente para presentar una estrategia de implementación.

### 3. Metodología para el desarrollo de sitios WWW

**Diseñar** un sistema de manejo de archivos que satisfaga adecuadamente las necesidades de la implementación.

**Crear** plantillas y componentes que puedan ser usados para implementar con rapidez y eficiencia el sitio.

**Escribir** documentos HTML que implementen el sitio tanto en forma manual, como utilizando herramientas de edición para HTML.

**Mantener** en operación los archivos HTML, considerando su actualización, funcionamiento y aspectos técnicos.

El proceso de implementación, entonces, es el de realizar un trabajo disponible y en uso.

#### 3.3.5 Promoción

Este proceso tiene como meta que el sitio que ha sido construido o modificado sea "visitado" por las personas a las que va dirigido el mensaje del mismo.

La función de un promotor es mantener al tanto de los objetivos y oferta que el sitio ofrece al público en general, y los usuarios constantes. El promotor debe tener destreza en el manejo de relaciones públicas, comunicación interpersonal y comunicación de masas.

Para el éxito de este proceso se debe tomar en cuenta los siguientes aspectos :

**La audiencia** : La audiencia es fundamental para el Web, ya que para ésta han sido definido y analizado cada uno de los procesos que intervinieron en la creación del sitio.

**Comercial o no comercial** : El Web se ha transformado en una comunidad de personas y no una colección inerte de máquinas y programas, por lo que si un promotor considera únicamente los factores tecnológicos se encontrará en un error. El Web se ha convertido en un medio excelente para la convivencia de las personas así como para el comercio a nivel mundial.

**Medios de publicidad** : Existen diversos medios de comunicación que son útiles para divulgar la existencia de un sitio. Los medios tradicionales por lo general se enfocan a grupos locales, en cambio, los medios electrónicos modernos tales como los boletines electrónicos, grupos de noticias, listas de discusión, sitios Web, entre otros, permiten llegar a un amplio público a nivel mundial.

### 3. Metodología para el desarrollo de sitios WWW

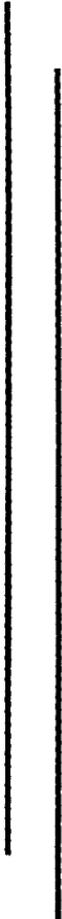
Con estos elementos, el promotor puede estructurar una estrategia de publicidad de acuerdo a sus recursos (monetarios, de infraestructura, etcétera), alcances y conveniencias particulares.

#### 3.3.6 Innovación

La innovación es un proceso dinámico y creativo, que no puede ser encapsulado en una serie de pasos establecidos. La innovación es un repertorio de habilidades en el monitoreo y entendimiento de las necesidades y desarrollos en las estructuras que forman el Web.

Los desarrolladores deben mantenerse informados sobre sitios similares a los suyos o de competidores. Si es factible, se puede considerar la colaboración en los sitios de la competencia, ya que permite llegar a un amplio público a nivel mundial tal que cada organización pueda enfocar su atención en elementos claves para cada una de ellas.

Una técnica importante para la innovación en el WWW es la mejora continua de los sitios que se tengan. Esto se logra conociendo las necesidades de los usuarios para implementar correcciones, mejorar la accesibilidad del sitio, crear programas y/o código reutilizable y hacer más interesante el sitio.



# CAPÍTULO 4

Aplicación de la  
metodología

## 4. Aplicación de la metodología

En el capítulo anterior se presentó y analizó cada uno de los elementos y procesos que integran la metodología de creación de sitios WWW. En el presente capítulo se utilizará para la puesta en funcionamiento del sitio del Departamento de Evaluación de Aplicaciones.

### 4.1 Antecedentes

El *Departamento de Evaluación de Aplicaciones* forma parte de la *Coordinación de Servicios de Red de la Dirección de Cómputo para la Administración Académica* (DCAA). Tiene por objetivo el de evaluar *software* y *hardware* a fin de determinar las cualidades que poseen y determinar el uso adecuado de estos, además, tiene bajo su administración sistemas de información internas (atención a equipos, servidores, Intranet, entre otros).

El departamento genera productos tales como sistemas, fichas técnicas, manuales, emisión de recomendaciones, que son de gran valor en la UNAM tanto para los desarrolladores de *software* y/o líderes de proyectos. Muchos de estos productos han sido elaborados por peticiones concretas, y por sus características pueden solucionar problemáticas similares en otras dependencias Universitarias; o bien, los líderes de proyectos no cuentan con referencias técnicas ni profesionales sobre algún producto en particular.

Toda esta información se da a conocer en tres áreas: Sistemas a cargo del departamento, Aplicaciones evaluadas por el departamento y Aplicaciones elaboradas por el departamento. Estas divisiones abarcan principalmente los resultados prácticos y de uso inmediato con los que cuenta el *Departamento de Evaluación de Aplicaciones*. Ante esta situación, y como parte de las tareas de difusión propias del departamento, se ha determinado la creación del sitio WWW del mismo.

Con estas premisas se utilizará la metodología para desarrollar el sitio del *Departamento de Evaluación de Aplicaciones*.

#### 4. Aplicación de la metodología

##### 4.2 Planeación

Como lo define la metodología, en este proceso se debe efectuar el análisis de los elementos que contendrá el sitio, así como el análisis correspondiente a la puesta en funcionamiento del mismo.

El planteamiento mismo del sitio define ya tres áreas de interés que se deberán de abordar, al igual que determina a quienes está dirigido el sitio : desarrolladores de software y líderes de proyectos de las diferentes dependencias Universitarias. Se agrega a estas áreas una más que servirá para obtener información de los visitantes que accesan el sitio, éste módulo se denomina "Libro de visitas".

La información que se presentará en el sitio ya ha sido recopilada y se dispone de ella en base a manuales, tutoriales, sitios en línea, la documentación propia de cada uno de los productos y la documentación que genera el departamento, entre otros.

El departamento cuenta además con la infraestructura adecuada para disponer de un sitio propio, así como de los profesionales que lo desarrollarán y que posteriormente lo administrarán.

El equipo que se ha destinado para la puesta del sitio es un servidor SUN SPARCstation 4, que cuenta con sistema operativo Solaris 2.4 (UNIX). Se utilizará el servidor de NCSA versión 1.5.2 de dominio público<sup>42</sup>.

El servidor que se levantó para el sitio tiene las siguientes características :

Dirección IP del servidor : 132.248.71.5 ó azteca.dgsc.a.unam.mx  
Puerto de funcionamiento : 8080  
URL de la página principal : <http://azteca.dgsc.a.unam.mx:8080/Tesis/Principal.html>

Una vez definido los elementos generales que regirán la creación del sitio, se puede pasar al proceso de análisis para continuar estableciendo los lineamientos que se seguirán.

---

<sup>42</sup> Para mayor información sobre la instalación, configuración y seguridad sobre el servidor NCSA, se sugiere al lector la visita de su sitio en : <http://hoohoon.ncsa.uiuc.edu/>

### 4.3 Análisis

La información que se presentará al usuario es muy variada y dependerá de los objetivos particulares de cada área (Sistemas a cargo del departamento, aplicaciones evaluadas por el departamento, aplicaciones elaboradas por el departamento y el módulo de consulta), pero deberá de existir una página de referencia la cual permita consultar las demás secciones del sitio, es decir, una página principal. La página principal, como se mencionó en la sección 3.3.2 del capítulo anterior, debe proporcionar al usuario un panorama general del sitio.

Aunque existe una página principal del sitio, es importante señalar que en cada área se definen características propias y que deben ser consideradas para su diseño particular. Por ejemplo, la organización de la sección *Sistemas a cargo del departamento* debe ser diferente a la que utiliza *Aplicaciones evaluadas por el departamento* y *Aplicaciones elaboradas por el departamento*. Por otro lado, el uso de formas<sup>43</sup> es obligatoria para la sección del "Libro de visitas".

Además, se necesita establecer un orden jerárquico de las páginas que se diseñen, es decir, que existan páginas intermedias que ayuden a localizar la información en el sitio. Esta jerarquía se integrará a partir de la página principal siguiendo los diferentes niveles que se formen en cada uno de los módulos. Siempre habrá alguna forma de ir al nivel inmediato superior.

Con base en los ejemplos expuestos en la Sección 3.3.2 del capítulo anterior los siguientes puntos se definen como útiles para este sitio :

- Elementos básicos de HTML que puedan ser interpretados por cualquier Navegador<sup>44</sup> :
  - Encabezados.
  - Estilos de letras (negritas, itálicas, subrayado, etc.).
  - Líneas horizontales.
  - Listas (numeradas, no numeradas y de definiciones).
- Elementos básicos de las formas<sup>45</sup> :
  - Botones.
  - Campos de texto.
- Dividir la información en niveles :
  - Establecer páginas intermedias.

<sup>43</sup> Sección de una página WWW para la captura de datos que proporciona el usuario. Ver Anexo B - CGI (Common Gateway Interface).

<sup>44</sup> Ver Anexo A - El lenguaje HTML.

<sup>45</sup> Ver Anexo B - CGI (Common Gateway Interface).

#### 4. Aplicación de la metodología

- Presentar la información en tópicos bien definidos.

No se desea el uso de imágenes como ligas a otras páginas, pero el uso del hipertexto se verá limitado por la información que se disponga, esto es, solo habrá ligas que estén asociadas a un documento o fuente de información que realmente exista.

Con estos planteamientos del análisis se procede entonces a definir un diseño propio.

#### **4.4 Diseño**

En el proceso de análisis, el equipo de trabajo evaluó diferentes propuestas, a través de un sondeo de otros sitios y su propia experiencia. Con estos elementos y otros que se han considerado pertinentes, ya se tiene la pauta para crear el sitio. Como se mencionó el proceso de diseño busca establecer una vista y una funcionalidad propia adecuada al sitio.

Para la página principal del sitio del Departamento de Evaluación de Aplicaciones, se ha contemplado el siguiente diseño :

La página principal será el "Índice" del sitio. Es decir, de ésta se parte a cualquier otro módulo o página. Debe presentar información general y concreta. Debe ser de fácil navegación para el visitante. Un esquema general de la página principal se muestra en la Figura 4.1.

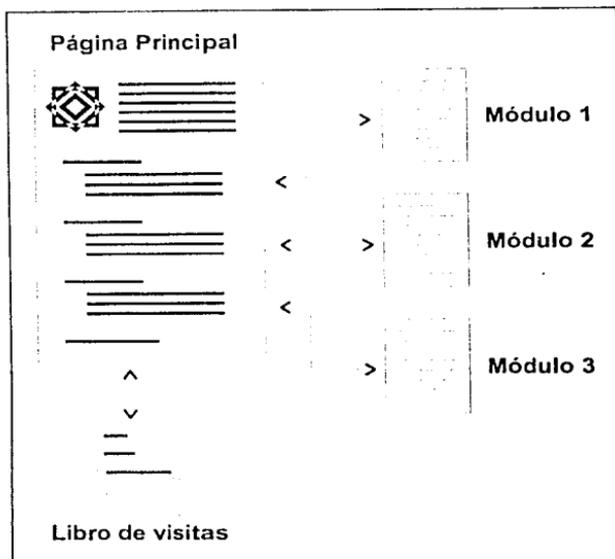


Figura 4. 1 Diseño de la página principal

A continuación se definen los contenidos de cada uno de los módulos que se han definido :

- Sistemas a cargo del departamento

Este módulo presentará una lista con los títulos de cada uno de los sistemas que tiene a su cargo el departamento para su correcto funcionamiento. Los elementos de dicha lista serán hipertexto a las páginas de información de cada sistema. Además esta lista debe tener un hipertexto a la página principal del sitio. La información que se presentará en las subsecuentes páginas es la siguiente : Nombre del sistema, responsable(s) y características sobresalientes. Las páginas de información deben tener además hipertexto a la lista y a la página principal. Ver Figura 4.2.

#### 4. Aplicación de la metodología

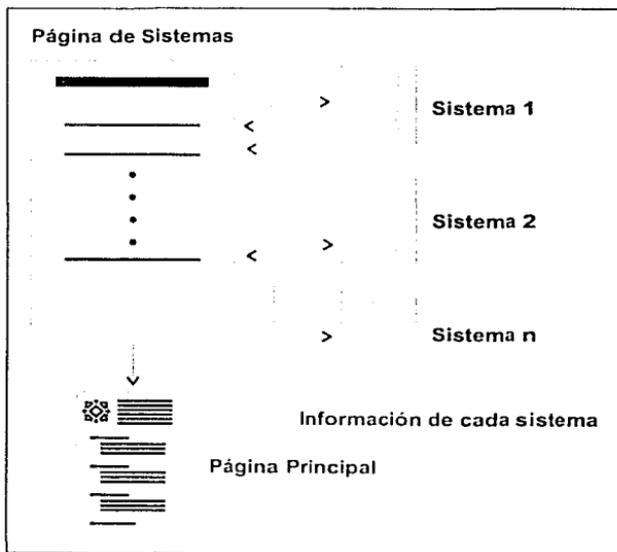


Figura 4. 2 Diseño para el módulo "Sistemas a cargo del departamento"

- Aplicaciones evaluadas por el departamento

Este módulo es similar al anterior, salvo que las páginas de información de los productos evaluados presentarán la siguiente información : Nombre del producto, versión, periodo de evaluación, tipo de licencia, página del fabricante, versiones disponibles, descripción general del producto y liga a una copia del documento evaluado. Ver Figura 4.3.

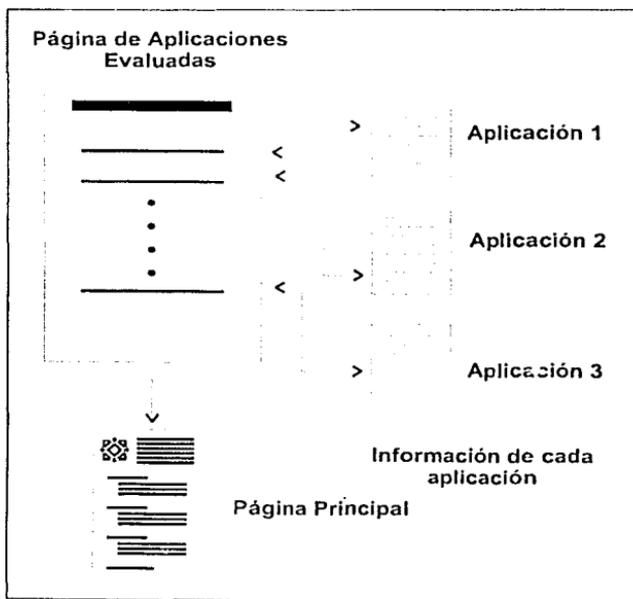


Figura 4.3 Diseño para el módulo "Aplicaciones evaluadas por el departamento"

- Aplicaciones elaboradas por el departamento

En este módulo se listarán los programas que se han creado, su funcionamiento y su diseño es similar al de los módulos anteriores. La información que se presenta es la siguiente: Nombre del producto, responsable, dirección electrónica del responsable, descripción general del producto.

- Libro de visitas

Este módulo difiere de los demás, ya que aquí el usuario puede registrar su visita al sitio. En este caso se necesita planificar un formulario de registro

ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA

#### 4. Aplicación de la metodología

de datos, que son : Nombre del visitante, dirección de correo electrónico, URL de su HomePage, ciudad donde reside, estado, país y sus comentarios. Ver Figura 4.4.

El diagrama muestra un formulario web con los siguientes elementos:

- Título**: Encabezado principal del formulario.
- Campos de entrada**:
  - Nombre
  - E-mail
  - Home Page
  - Ciudad, Edo., País (campos separados)
  - Comentarios
- Botones**: "Enviar" y "Limpiar forma".
- Registro de visitas**: Una tabla con líneas horizontales que representan los datos de los visitantes.
- Menú de navegación**: Un icono de cursor y un enlace etiquetado como "Página Principal".

Figura 4. 4 Diseño para "Libro de Visitas"

Una vez que se ha establecido el diseño cada una de las partes que integran un sitio, falta ahora crearlo.

#### 4.5 Implementación

La página principal así como los tres primeros módulos que forman el sitio ya pueden ser generados con el código de HTML presentado. Las páginas obviamente se apegan al diseño que ya se había especificado.

Como se había indicado, la presentación final que el usuario tendrá se deberá en parte a la disposición que de los elementos se dispuso, y en otra medida al Navegador que el usuario utilice. A continuación se presenta la página principal del Departamento de Evaluación de Aplicaciones desplegada por diferentes Navegadores.

- Con Microsoft Explorer :



Figura 4. 5 Página principal (1)

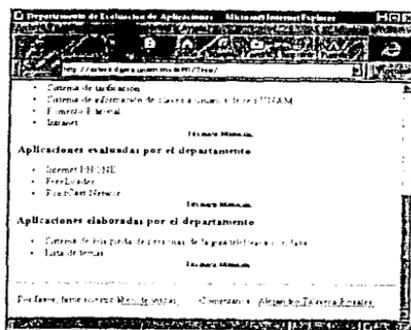


Figura 4. 6 Página principal (2)

#### 4. Aplicación de la metodología

Con NetScape :



Figura 4. 7 Página principal (1)



Figura 4. 8 Página principal (2)

Como vemos cada uno de ellos tienen ciertas diferencias con respecto de los demás. Algunos Navegadores tienen incluidos etiquetas que otros no tienen, o bien, no lo presentan de forma idéntica en los demás.

Ahora se mostrarán los demás elementos del sitio una vez que han sido creadas las páginas HTML.

#### Sistemas a cargo del departamento

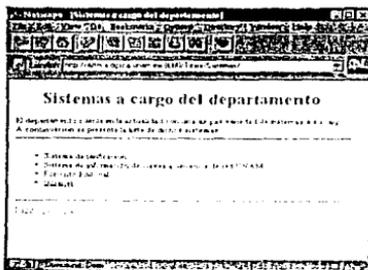


Figura 4. 9 Página para el módulo "Sistemas a cargo del departamento"

#### Aplicaciones evaluadas por el departamento



Figura 4. 10 Página para el módulo "Aplicaciones evaluadas por el departamento"

#### 4. Aplicación de la metodología

##### Aplicaciones elaboradas por el departamento

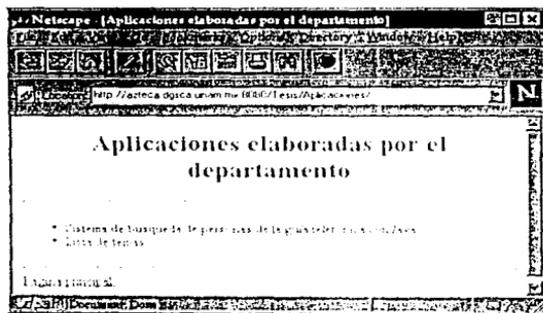


Figura 4. 11 Página para el módulo "Aplicaciones elaboradas por el departamento"

Para el Libro de Visitantes que ha especificado, se necesitará otro tipo de elementos, ya que el usuario va a dar información que será procesada y puesta en un documento HTML para su consulta posterior. Para este módulo en especial, se necesitará el empleo de programas CGI's.

El CGI es un estándar que define una serie de variables y convenciones para crear programas que sirvan de interfaz<sup>46</sup> entre el servicio de WWW y una aplicación en especial. Los programas que se apegan a dicho estándar son normalmente conocidos como CGI<sup>47</sup>.

Volviendo al caso del Libro de Visitantes, para el diseño se tendrá que utilizar algunos de estos elementos. El programa CGI de esta opción recupera la información que ofrece el usuario, y actualiza un archivo para mantener el registro de visitas. El programa CGI que se va a utilizar tiene el siguiente URL : <http://azteca.dgsca.unam.mx:8080/cgi-bin/guestbook.pl>.

El código HTML que genera la forma, según la especificación que se ha definido, es el siguiente :

<sup>46</sup> Un mecanismo que permite la comunicación entre dos protocolos o programas diferentes.

<sup>47</sup> Para mayor explicación de los CGI's ver el Anexo B.

```

<HTML>
<HEAD>
<TITLE>
  Firme el Libro de visitas
</TITLE>
</HEAD>
<BODY BGCOLOR="#FFFFFF">
  <CENTER>
    <h1>Firme el Libro de visitas</h1>
  </CENTER>

  Llene los espacios en blanco del formulario para poner sus datos en
  nuestro libro de visitas. Los datos que al menos se deben llenar son
  los comentarios y su nombre. Gracias!

  <HR>

  <FORM METHOD=POST ACTION="http://azteca.dgsca.unam.mx:8080/cgi-bin/guestbook.pl">
  Su nombre : <INPUT TYPE=TEXT NAME=realname SIZE=30><BR>
  E-Mail : <INPUT TYPE=TEXT NAME=username SIZE=40><BR>
  URL de su página en WWW : <INPUT TYPE=TEXT NAME=url SIZE=50><BR>
  Ciudad : <INPUT TYPE=TEXT NAME=city SIZE=15>,
  Estado : <INPUT TYPE=TEXT NAME=state SIZE=2>
  Países : <INPUT TYPE=TEXT VALUE="MEXICO" NAME=country SIZE=15><P>
  Comentarios :<BR>
  <TEXTAREA NAME=comments COLS=60 ROWS=4></TEXTAREA><P>
  <INPUT TYPE=SUBMIT VALUE="Enviar datos"> <INPUT TYPE=RESET VALUE="Borrar forma">
</FORM>
  <HR>
  <A HREF="http://azteca.dgsca.unam.mx:8080/Tesis/Guest/guestbook.html">
  Ver el libro de visitas</A><BR>
</BODY>
</HTML>

```

El programa guestbook.pl desarrolla las siguientes acciones :

1. Obtiene las parejas de nombre - valor de cada elemento de la forma.
2. Actualiza el documento de visitantes con los datos anteriores.
3. Presenta un mensaje notificando la inserción de información.

A continuación se presentan la ventana que presenta el Libro de visitantes.

The image shows a screenshot of a web browser window. The title bar indicates the browser is Netscape 3.0. The address bar shows a URL starting with 'http://www...'. The main content area has a heading 'Firme el Libro de visitas' (Sign the Guest Book). Below the heading is a short paragraph of text. The form contains several input fields: 'Nombre' (Name), 'E-Mail', 'URL de su página en WWW', 'Ciudad', 'País', and 'E-mail (Opcional)'. There is also a large text area for a message. At the bottom of the form are two buttons: 'Enviar Registro' and 'Diferílo más tarde'. The browser's status bar at the bottom shows the current page is 'http://www...'. The browser's menu bar includes File, Edit, View, Bookmarks, Options, Directory, Window, and Help.

Figura 4. 12 Página para el módulo "Libro de visitantes"

#### 4.6 Promoción

En el caso del sitio del Departamento de Evaluación de Aplicaciones la publicidad será la siguiente :

- Se dará a conocer en los foros de discusión de la UNAM.
- Emisión de oficios por parte del departamento anunciándolo.
- En las presentaciones de cómputo en donde participe el departamento.

#### 4.7 Innovación

Este es uno de los puntos, como se mencionó, más difíciles de cumplir. Algunas de las características más evidentes en *Internet* es la distribución y uso de nuevas tecnologías, ya no se diga el innovar. Cada día más sitios se ponen en funcionamiento y se hace más importante el encontrar técnicas y diseños que mejoren un sitio.

La innovación, como se refirió en la presentación de la metodología, consiste en encontrar elementos que favorezcan el sitio y lo hagan más atractivo y práctico. Esto involucra mejorar las formas de navegación dentro del sitio, depurar los programas CGI a fin de utilizar óptimamente los recursos del servidor, y el monitoreo de nuevas técnicas para integrarlas a los sitios que se mantienen o se crean.

Un ejemplo sobre la búsqueda de la innovación se puede apreciar en la sección de "Productos elaborados por el departamento", particularmente los desarrollos con el lenguaje de programación Java.

Java presenta muchas características que otros lenguajes carecen. Algunas de estas características se listan a continuación:

- Es un lenguaje orientado a objetos.
- Es independiente a la plataforma.
- Tiene una amplia variedad de bibliotecas.
- Se pueden generar programas que sean ejecutados utilizando recursos de máquinas remotas denominados *applets*.
- Manejo independiente de aplicaciones gráficas.

Este lenguaje fue desarrollado por SUN Microsystems en el año de 1990. No fue hasta el año de 1993 que Java toma mayor importancia, cuando se desarrolla un Navegador de WWW denominado "Hot-Java" hecho con este lenguaje. Aunque carecía de una amplia variedad de elementos que los demás tenían (tales como la interpretación de muchas etiquetas de HTML o manejo de correo electrónico), Hot-Java presenta una característica única : es capaz de ejecutar en la máquina remota un programa, aunque debía de haber sido hecho con Java. En cambio a diferencia de los CGIs tradicionales utilizan recursos de la máquina servidora. En la actualidad existen tres Navegadores que permiten la interpretación de programas Java : Hot-Java, Netscape (versiones 2 o posteriores) y Microsoft Explorer, en plataformas UNIX, Macintosh y PC (con Windows 95 o Windows NT).

Los applets (programas que corren en una página Web), utilizan recursos del equipo que los consulta, con ello se puede pensar en la distribución de procesos en varios equipos.

Un ejemplo del empleo de esta tecnología es el producto "Lista de temas".

El planteamiento que origina la programación del applet, surge de las siguientes características :

Se tiene una lista de temas, cada uno asociado a un URL específico. Se desea hacer una búsqueda sobre los temas, de tal manera que cuando se escriba la

#### 4. Aplicación de la metodología

palabra de búsqueda se presente la mejor ocurrencia, en este caso, el título que comience con el texto que se va escribiendo.

Una solución sería facilitar al visitante un documento HTML con la lista de temas aprovechando el hipertexto. El problema con esta solución sería la actualización del archivo, ya que se tendría que respetar el formato HTML definido desde el inicio. Un cambio en la presentación sería un cambio general a todo el documento.

Otra solución a este problema, es la de crear un CGI, y su correspondiente forma, para solicitar al usuario dicha palabra clave. Entonces el CGI efectúa la búsqueda de la palabra en el archivo que tuviese la lista de temas. Dicho archivo tendría un formato definido, ajeno al HTML y el CGI daría formato a la información para desplegar un documento HTML.

La solución que se dio, con las características de Java, fue la de crear una estructura de búsqueda a la manera de las ayudas del sistema Windows. Los menús de ayuda a los que se hace referencia tienen las siguientes secciones :

El diagrama muestra una interfaz de usuario con tres secciones numeradas:

- Una caja de texto rectangular con el número 1 en el centro.
- Una lista de temas de ayuda, representada por una barra horizontal con el número 2 en el centro.
- Una barra de navegación con un botón rectangular que contiene el número 3.

1. Caja de texto donde el usuario introduce una palabra clave.
2. Lista de temas de ayuda. La cual cambia de posición, según lo que el usuario este escribiendo (no se necesita oprimir la tecla ENTER para efectuar una búsqueda).
3. Una vez encontrado el tema, con activar a la tecla ENTER o dar clic en un botón se va al tema señalado.

Esta solución tiene muchas ventajas :

- No se puede realizar con un programa CGI, ya que es imposible crear uno que haga exactamente lo mismo.
- No se necesita realizar el programa en el servidor.

#### 4. Aplicación de la metodología

- No se utilizan recursos del servidor (sólo se requiere enviar el applet al Navegador).
- El servidor solo tiene que tener una versión del programa (no existe versión para PC, UNIX, etc.)
- Cuando se actualice la información de los temas o de algún módulo del programa, los usuarios utilizarán la nueva versión.

Actualmente el applet tiene varios usos, de los que destacan los siguientes: permite a los visitantes un medio de navegación más dinámico, se pueden verificar más rápidamente URLs de otros sistemas, la referencia al archivo que contiene los temas y URLs se hace en el documento HTML y no en el programa. En la Figura 4.13 se ve el applet funcionando en una página WWW.



Figura 4.13 Lista de temas (Applet)

# CONCLUSIONES

## Conclusiones

Hoy en día, la competencia por captar un mayor número de visitantes a los diferentes sitios en la Internet es muy grande, surgiendo la necesidad de que los equipos de desarrollo cuenten con las herramientas y los conocimientos necesarios para enfrentar y solucionar problemas derivados de este tipo de desarrollo.

En el presente trabajo se presentó una metodología de desarrollo para la elaboración de sitios *Web*, así como su aplicación en un caso real: el sitio del Departamento de Evaluación de Aplicaciones, de la DGSCA (Dirección General de Servicios de Cómputo Académico).

Como se planteó al inicio de este trabajo, ninguna metodología garantiza un producto libre de errores, pero sí una guía ordenada y planificada de trabajo. Se ofrece además una explicación sobre los servicios básicos que ofrece la *Internet*, así como las bases teóricas y técnicas sobre las cuales se sustentan. Dicha explicación proporciona a los futuros usuarios las bases necesarias y suficientes para un mejor aprovechamiento de los mismos.

Los servicios aquí presentados, ofrecen al lector un panorama general del uso e impacto que actualmente tienen en la Internet, así como los posibles alcances que ésta pudiera tener.

Por otra parte, se ofrece una guía con los elementos básicos para que el lector pueda generar documentos WWW. Por lo que, al utilizar apropiadamente la metodología, estos puedan llegar a ser un buen y atractivo sitio *Web*. Sin embargo los desarrolladores pueden enriquecerla apoyándose en su experiencia, objetivos particulares y las capacidades técnicas y humanas con que cuentan.

Aunado a la anterior, se proporciona una segunda guía en la que se dan a conocer las bases teóricas para la generación de programas denominados CGIs, los cuales establecen una iteración dinámica entre el usuario final y los recursos del sistema.

Finalmente, podemos decir que existen muchas expectativas sobre los servicios en *Internet*. Muchos de ellos dependen de los desarrollos actuales y a mediano plazo, de los cuales se presentan los más importantes:

- En la parte de comunicaciones se necesita del desarrollo de infraestructura más rápida y versátil, que permita el manejo eficiente de la demanda de información, que existe ya en la Internet, así como su adaptación a nuevos sistemas. Se tienen, por ejemplo, planes de establecer servicios de video vía redes de

## Conclusiones

computadoras, lo cual involucra redes que permitan el paso de una gran cantidad de información de audio y video.

- La seguridad de la información que se mueve en *Internet*, es un factor de gran relevancia. Se necesitan nuevos sistemas de protección de la información ya que cada vez más información de diferente índole se mueve por las redes de computadoras. En la medida que se mejore la seguridad, se darán más oportunidades al comercio por la *Internet*. Esto acercará a los diferentes proveedores de servicios a sus posibles clientes potenciales, a nivel mundial.
- El procesamiento en paralelo de procesos complejos permitirá la creación de nuevas bases de nuevos servicios, así como la utilización eficiente de los equipos conectados a red.
- Otro aspecto importante será el de establecer estándares para la presentación de información. Actualmente, la búsqueda de información es un problema muy serio, pues como no existe un esquema para la presentación de información los métodos actuales resultan ser muy complicados de desarrollar.

# ANEXO A

## El lenguaje HTML

## El lenguaje HTML

El HTML (*HiperText Markup Language*), que traducido literalmente sería "Lenguaje de Marcas de Hipertexto", es el lenguaje para la generación de páginas Web. Fue desarrollado inicialmente por Tim Bernes-Lee en el CERN en el año de 1989. El HTML se origina como un formato para que los investigadores y científicos pudieran utilizar diversas computadoras (Pc, Macintosh, Unix, etc.), compartiendo información por red; para ello se requirió la "independencia de plataforma", que implica que un documento se puede desplegar de igual modo en computadoras con diferentes elementos (tipos de letras, gráficas y color). Otra característica es el "Hipertexto", que significa que cualquier palabra o frase de algún documento pueda hacer referencia a otro documento, para que permita hacer más fácil la navegación entre los numerosos documentos que haya en el sistema.

Berner-Lee decidió utilizar como patrón el SGML (*Standar Generator Markup Language*) "Lenguaje estándar generado por marcas", establecido por la norma ISO 8879:1976.

Este estándar internacional es independiente de la plataforma debido a que se enfoca a codificar la estructura semántica de un documento, no necesariamente por la apariencia que tenga. De esta forma, por ejemplo, el título de un capítulo quedará etiquetado con algo como "Título de capítulo", en vez de "Arial 12pt centrado". Cada programa lector define la apariencia del título de los capítulos de modo que sea útil en la computadora que se utiliza, y cualquier texto con formato se despliega en forma apropiada.

Una ventaja del SGML es su flexibilidad, ya que como no es propiamente un formato, sino una especificación para definir otros; los usuarios pueden crear nuevos formatos para codificar toda la estructura de determinados tipos de documentos (Por ejemplo, manuales técnicos, directorios telefónicos y documentos legales), y cualquier *software* habilitado para el SGML puede entenderla, con sólo leer la primera definición. A partir de dicho estándar se han creado una gran cantidad de lenguajes conocidos como DTD (*Document Type Definition*), tanto para documentos comunes como para documentos muy especializados. El HTML, en este contexto, es sólo un DTD o *aplicación* de SGML.

Cuando NCSA construyó por primera vez *Mosaic* a principios de 1993, agregó características propias al HTML, incluyendo gráficas en línea. En mayo de 1994 parecía que HTML crecía sin control, ya que los fabricantes de Navegadores agregaban características propias, muchas veces incompatibles al estándar SGML. En la primera conferencia sobre WWW en Ginebra Suiza, se organizó un

## Anexo A - El lenguaje HTML

grupo de trabajo en HTML, cuya tarea era la de formalizar el lenguaje. Conforme se utilizaba, dentro de un DTD conocido como HTML Nivel 2 (El Nivel 1 fue definido como fue diseñado originalmente por Berners-Lee). Una vez que se estandarizó, se pudo extender de una manera segura a los futuros niveles, e incluso aprovechar las posibilidades del SGML original y su estructura formal.

Un ejemplo sencillo de una página generada con HTML, es la siguiente:



Figura A.1. Página elaborada con HTML.

el código del documento es el siguiente:

```

<HTML>
<HEAD>
<TITLE>
  P&aacute;gina de ejemplo
</TITLE>
</HEAD>
<BODY>
  <H1> P&aacute;gina principal </H1>
  La presente p&aacute;gina ejemplifica el uso del lenguaje
  HTML, as&iacute; como el uso de elementos b&aacutesicos, tales
  como saltos de l&iacut;nea, im&aacute;genes e hipertexto.
  <HR>
  <H2> Saltos de l&iacut;nea </H2>
  Este texto presenta saltos de l&iacut;nea al final de cada una.<BR>
  Para lograr esto se utiliza<BR>
  la etiqueta &lt;BR>.

  <H2> Im&aacute;genes</H2>
  <IMG SRC="perro.gif">

  <H2> Hipertexto </H2>
  P&aacute;gina de la
  <A HREF="http://www.unam.mx/">
  UNAM
  </A>.
</BODY>
</HTML>

```

Como se mencionó, el HTML codifica sólo la estructura del documento. La mayor parte de la "apariciencia" del documento, como los estilos de escritura, el color y el tamaño de la ventana, se encuentran bajo el control del Navegador que lo despliega. Un documento HTML no debe tener ningún formato, es decir, debe ser guardado en código ASCII.

Como se observa en el listado anterior, un documento HTML consta de dos tipos de contenidos : el texto del documento y los códigos ó etiquetas. Éstas últimas son cadenas de textos encerradas por un signo de menor que (<) y otro de mayor que (>), como <HTML> de la primera línea.

La sintaxis general de una etiqueta es la siguiente :

## Anexo A - El lenguaje HTML

```
<nombre_de_la_etiqueta atributo1=valor1 atributo2=valor2 ... atributon=valorn>  
  Texto y/o elementos afectados por la etiqueta.  
</nombre_de_la_etiqueta >
```

El nombre de la etiqueta es el tipo de texto que define la etiqueta; los atributos dan información adicional acerca de cómo se presentará el elemento; algunas etiquetas carecen de atributos. Cabe mencionar que el nombre de la etiqueta y el de los atributos pueden ser escritos en mayúsculas y minúsculas. En cambio, los valores de cada atributo pueden serlo o no de acuerdo a la plataforma en donde se almacene el documento.

La última etiqueta, que comienza con una diagonal (/), indica el límite dentro del documento que dicho código afecta. Algunas etiquetas carecen de ésta última. Por ejemplo, en la tercera línea del listado anterior, las etiquetas <TITLE> y </TITLE> definen al texto contenido como entre ellas como un título. En cambio la etiqueta <IMG SRC=... > sólo consta de una y no altera texto alguno; en su lugar inserta una imagen en el documento.

Las etiquetas y el texto que enmarcan se combinan para formar elementos. Cada elemento representa un objeto en el documento, como un encabezado, un párrafo o una imagen. Un elemento consta de un grupo o dos etiquetas y por lo general de algún texto asociado.

A continuación se presentan las sintaxis de algunas etiquetas básicas para la elaboración de documentos.

### Documento HTML

```
<HTML>  
  El documento  
</HTML>
```

Esta etiqueta indica que el documento es HTML y abarca todo el archivo. Actualmente su uso es optativo.

### Encabezado del documento

```
<HEAD>  
  Encabezado del documento  
</HEAD>
```

Esta etiqueta se utiliza para indicar la parte de cabecera de un documento HTML. En la cabecera se coloca el título del documento.

**Título del documento**

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

Se utiliza para indicar el texto que va a servir como título al documento. Va entre las etiquetas <HEAD> y </HEAD>.

**Cuerpo del documento**

```
<BODY>  
  El documento  
</BODY>
```

Este elemento contiene otros elementos que representan el cuerpo de texto del documento, por lo general toda la extensión del archivo.

En si un documento HTML se ve esquematizado de la siguiente forma por estos tres elementos :

```
<HTML>  
  <HEAD>  
    <TITLE>  
      Título del documento  
    </TITLE>  
  </HEAD>  
  <BODY>  
    Texto y/u otros elementos del documento  
  </BODY>  
</HTML>
```

Los siguientes elementos que se listan, deben estar contenidos entre las etiquetas <BODY> y </BODY>.

## Anexo A - El lenguaje HTML

### Salto de línea

<BR>

Con esta etiqueta se fuerza un salto de línea, es decir, el texto que va después de esta etiqueta comenzará en la siguiente línea. Esta etiqueta utiliza </BR> para su funcionamiento.

### Párrafos

<P>

Con una sola etiqueta se indica un salto de párrafo, es decir, deja una línea en blanco antes del siguiente bloque de texto. Es como si se utilizaran dos etiquetas <BR>, pero con la diferencia que no son acumulativas.

La etiqueta final (</P>) se utiliza para indicar un párrafo. Si se utiliza junto con el parámetro ALIGN, que indica la alineación del texto del párrafo. Este atributo puede tomar cualquiera de los siguientes tres valores de alineamiento : "RIGHT" "LEFT" o "CENTER" para alineamiento a la derecha, izquierda o centrada respectivamente.

<P ALIGN=valor>  
    *Texto del párrafo*  
</P>

### Línea

<HR>

Este elemento se utiliza para desplegar una línea horizontal en el documento. No requiere etiqueta final.

### Estilos

Los siguientes elementos se utilizan para cambiar el tipo de letra que se haya enmarcado :

<B> texto </B>      Texto en negritas.  
<I> texto </I>      Texto en cursiva.  
<U> texto </U>      Texto subrayado<sup>48</sup>.

<sup>48</sup> No es ampliamente soportado por la mayoría de los Navegadores.

Se pueden utilizar estos estilos para generar combinaciones de ellos. Por ejemplo un texto en negritas con cursiva queda de la siguiente forma :

`<B><i>` Texto en negritas y cursiva `</i></B>`

### **Encabezados**

`<H#>` Texto del título `</H#>`

HTML define seis estilos de encabezados, numerados del 1 al 6. Dicha numeración representa la "importancia" del título. Estos encabezados no son afectados por otros atributos de estilo

### **Dirección de correo**

Es un estilo definido para la presentación de direcciones de correo electrónico. Se presenta en un tamaño más pequeño de letra y cursiva.

`<ADDRESS>` Dirección `</ADDRESS>`

### **Texto preformateado**

`<PRE>`

*Texto preformateado*

`</PRE>`

El HTML ignora espacios extra, saltos de línea, así como tabuladores en el documento. Algunos documentos, como listados de programas, tablas o poesía, entre otros, serían difíciles de codificar. Para desplegarlos con el formato que fueron creados se utiliza la siguiente etiqueta.

### **Listas**

Existen dos tipos de listas las numeradas y las no numeradas.

Las listas numeradas son útiles para presentar temarios, índices, etc. Todas las líneas que la forman se encuentran numeradas y no es necesario numerarlas ya que automáticamente se efectúa la numeración.

## Anexo A - El lenguaje HTML

```
<OL>
  <LI> Elemento 1 de la lista
  <LI> Elemento 2 de la lista
  <LI> Elemento 3 de la lista
  .....
  <LI> Elemento n de la lista
</OL>
```

Las listas no numeradas son muy parecidas a las anteriores, salvo que en lugar de números se presentan símbolos (puntos, diamantes, etc) para cada nivel.

```
<UL>
  <LI> Elemento 1 de la lista
  <LI> Elemento 2 de la lista
  <LI> Elemento 3 de la lista
  .....
  <LI> Elemento n de la lista
</UL>
```

Se pueden crear niveles, ya sea definiendo otra lista del mismo tipo o bien combinando estas listas.

```
<OL>
  <LI> Elemento 1 de la lista
  <UL>
    <LI> Sub-elemento 1
    <LI> Sub-elemento 2
    <LI> Sub-elemento 3
  </UL>
  <LI> Elemento 2 de la lista
  <LI> Elemento 3 de la lista
  .....
  <LI> Elemento n de la lista
</OL>
```

### Lista de definiciones

Estos elementos son útiles para presentar una sección de definiciones en un documento. No se numeran o son marcados con símbolos, sólo varía la sangría de los términos a definir y las definiciones en sí.

```
<DL>
  <DT> Término1 a definir
  <DD> Definición del término1
  <DT> Término2 a definir
  <DD> Definición del término2
  <DT> Término2 a definir
  <DD> Definición del término3
</DL>
```

### Imágenes

```
<IMG SRC="URL" ALING=TOP/MIDDLE/BOTTOM>
```

Con ésta etiqueta se colocan imágenes en un documento. La imagen debe de estar ya creada y se hace referencia con el atributo SRC, cuyo valor será el URL del archivo que contiene la imagen que se desea insertar en el texto. Otro atributo importante es el ALT cuyo valor será un texto alternativo por si el Navegador que consulta el documento no puede desplegar esa imagen. El atributo ALIGN indica el tipo de alineamiento que tendrá el texto que le siga. Los formatos de imágenes usados en el WWW son dos : GIF (*Graphics Interchange Format*) y JPEG.

### Hipertexto

Estas etiquetas están definidas para marcar texto que está asociado a un documento. Es texto marcado con esta etiqueta se presenta en otro color (generalmente en azul y subrayado), diferente al texto normal (Generalmente en negro). El valor de la opción HREF es el URL del documento con el que se asocia a dicho texto.

```
<A HREF="URL"> Texto marcado </A>
```

El Hipertexto es la asociación de un texto a otro. Sin embargo en WWW dicha asociación es más amplia, ya que se puede asociar un texto a una imagen, un texto a un sonido, una imagen a un texto, etc. Este concepto se conoce con el nombre de Hipermedia. Para lograr, por ejemplo, que una imagen esté asociado a un documento o a otro elemento multimedia, su declaración debe estar en medio de estas etiquetas.

```
<A HREF="URL"> Elemento multimedia </A>
```

Por ejemplo :

## Anexo A - El lenguaje HTML

```
<A HREF="http://www.unam.mx">  
  <IMG SRC="escudo.gif">  
</A>
```

Aquí la imagen cuyo archivo es "escudo.gif" será una "liga" a la página de la UNAM.

Cabe mencionar que tanto en el caso del hipertexto y de las imágenes, las fuente ó documentos pueden estar en otro equipo en la Internet diferente al que alberga el documento.

Así mismo, es importante indicar que si una etiqueta necesita su correspondiente etiqueta final todo el texto que le siga será afectado por esta. Así por ejemplo, Si se usa la etiqueta <B> y no se pone la etiqueta </B>, todo el texto que le siga presentará en negritas. En el caso de las etiquetas <A> y </A> todo el texto que siga a <A> será marcado como hipertexto.

Por otra parte, una etiqueta mal escrita, o que el Navegador utilizado no tiene declarada, no desplegará el efecto que se pretendía. Por ejemplo, Las etiquetas <TABLE></TABLE> están contempladas para el Nivel de la especificación del HTML. Los Navegadores diseñados para el Nivel 2 no presentaran la tabla que se había especificado y la información perderá el orden que se había planificado.

Las etiquetas aquí descritas son unas cuantas de las que actualmente existen. Para mayor referencia se sugiere al lector los siguientes documentos en línea :

- RGB Lab : <http://w3.cs.com.uy/u/rgblab/>
- NCSA--A Beginner's Guide to HTML Home Page :  
<http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html>
- Yale C/AIM Web Style Guide : <http://info.med.yale.edu/caim/manual/>

# ANEXO B

CGI  
*(Common Gateway  
Interface)*

## CGI (*Common Gateway Interface*)

El CGI es un conjunto de reglas para la conexión de aplicaciones externas con la información que tienen los servidores, tal como los servidores HTTP o *Web*. Un documento HTML proporcionados por un servidor Web es estático, lo cual significa que es un archivo de texto que no cambia. Un programa CGI, por otra parte, se ejecuta en tiempo real, por lo que puede proporcionar información dinámica.

Por ejemplo, si se desea "conectar" una base de datos de una plataforma UNIX al *World Wide Web*, para permitir su consulta a usuarios de todo el mundo, se necesita crear un programa CGI que el servidor *Web* ejecutará para transmitir información al controlador de la base de datos, recibir los resultados de la consulta y desplegarlos al cliente.

El ejemplo de la base de datos es una idea sencilla, pero se requiere de mucho tiempo para implementarla. Pero esto no es realmente una limitante para que se pueda realizar la conexión al *Web*, la única cosa que se necesita tener en mente es que cualquier cosa que un programa CGI haga, no debe ser un proceso demasiado largo. De otra forma, el usuario que ha solicitado una consulta estará en su Navegador esperando a que algo ocurra.

Como un programa CGI es ejecutable, esto equivale básicamente a permitir que cualquiera corra el programa en el sistema que lo almacena, lo que no es seguro de hacer. Sin embargo, existen medidas de seguridad que se deben ser implementadas al usar programas CGI. El único efecto que el usuario probablemente notará, será el hecho de que los programas CGI necesitan residir en un directorio especial, para que el servidor *Web* lo sepa antes de desplegarlo al Navegador. Este directorio está usualmente bajo el control del administrador del servidor *Web*, restringiendo al usuario común la creación de programas CGI. Hay otras formas para permitir el acceso a los scripts de CGI, pero esto depende de las políticas de administración establecidas.

Si se tiene la versión del servidor NCSA HTTPd, existe un directorio denominado "/cgi-bin". Este es el directorio donde todos los programas CGI residen actualmente. Un programa CGI puede ser escrito en cualquier lenguaje que le permita ser ejecutado en el sistema, tal como :

- C/C++
- Fortran
- PERL

## Anexo B - CGI (Common Gateway Interface)

- TCL
- SHELL (para cualquier versión de UNIX)
- Visual Basic
- Delphi
- AppletScript

Si se cuenta con un lenguaje de programación como C o Fortran, se necesita compilar los programas antes de que ser ejecutados. En el caso del servidor de NCSA HTTPd el directorio "/cgi-src" contiene los códigos fuente de algunos programas CGI que están en el directorio "/cgi-bin". Sin embargo, si se tiene en el sistema un lenguaje script, tal como PERL, TCL o Shell de UNIX, el script en sí mismo debe residir únicamente en el directorio "/cgi-bin", ya que no existe un código fuente asociado. Muchos programadores prefieren escribir scripts CGI en vez de programas, ya que son más sencillos de depurar, modificar y mantener que los típicos programas compilados.

Para hacer llegar información a un programa CGI, es necesario que el usuario utilice una estructura de forma definida por el lenguaje HTML. Los elementos de una forma pueden variar dependiendo de la información que se desea el usuario proporcione<sup>42</sup>. A continuación se presenta el funcionamiento de un programa CGI con un servidor *Web* (ver Figura C.1):

1. Luego que el usuario ha llenado un formulario de información (forma) en un Navegador, él o ella activa un botón que, es un vínculo hacia la aplicación CGI en el servidor.
2. El Navegador utiliza los parámetros de la definición de la forma para enviar los datos al servidor.
3. Los datos viajan a través de la *Internet* y llegan al servidor, pasando entonces hacia el programa CGI.
4. Se analizan los datos de la forma, según la petición del cliente, se procesan y el programa CGI "escribe" la respuesta en formato HTML que se enviará de regreso al cliente. La especificación del CGI establece que el servidor Web deberá "leer" el dispositivo definido como salida estándar del programa CGI.
5. El servidor agrega la información necesaria del encabezado HTTP y envía la salida del programa CGI a través de la Internet como un archivo HTML, que el Navegador recibe.
6. El Navegador interpreta el código HTML y despliega los resultados en pantalla para el usuario. Este archivo generalmente deberá contener al menos, alguna notificación de que los datos fueron procesados por el servidor.

---

<sup>42</sup> Consultar el Anexo A relacionado con la declaración de formas en HTML.

## Anexo B - CGI (Common Gateway Interface)

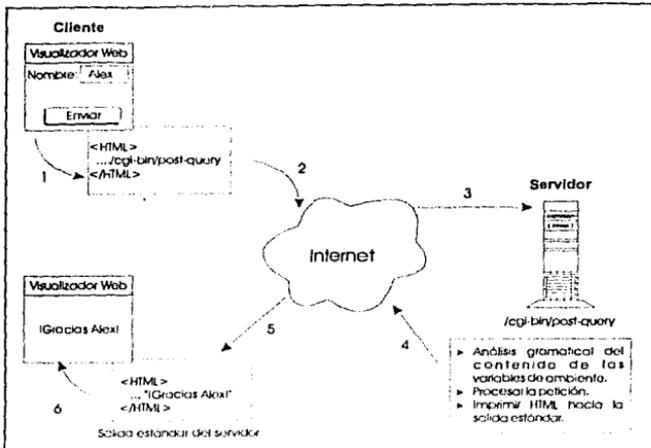


Figura C. 1 Funcionamiento del estándar CGI

Como se puede apreciar, los CGI utilizan recursos del servidor que los proporciona. Es decir, el servidor debe contar con los compiladores adecuados, los controladores de bases de datos (si se utilizan bases de datos), etcétera.

Lo que el visitante ve, lo proporciona el HTML, que define etiquetas para el manejo de Formas.

Una de las principales características de las formas es que permiten al usuario interactuar con los CGI, dando lugar a sitios "dinámicos". Las formas permiten al visitante proporcionar datos, que serán luego procesados por los CGI. Con estas formas se puede solicitar información para la consulta en una base de datos, solicitar inscripciones, anexar información, etc.

Para presentar un formulario en una página, el HTML utiliza las etiquetas `<FORM>` y `</FORM>`. Entre estas etiquetas puede ir cualquier elemento HTML que se ha presentado, excepto otra forma. La sintaxis de las etiquetas es la siguiente :

```
<FORM ACTION="URL_CGI" METHOD="GET|POST">
  Cuerpo de la forma
</FORM>
```

## Anexo B - CGI (Common Gateway Interface)

Los atributos que se pueden usar son los siguientes :

- A. **ACTION** : especifica el URL del programa que procesará la información que el usuario suministró.
- B. **METHOD** : es el método usado para enviar los datos obtenidos al programa que los procesará. El método que se usa depende de la forma en que trabaja el servidor. Las opciones válidas son los siguientes :
  1. **GET** : Este es el método pre-definido e indica que la información obtenida, será agregada al URL del programa.
  2. **POST** : Este método indica que la información obtenida en la forma será enviada al servidor en el cuerpo de la transmisión, en vez que en el URL
- C. **ENCTYPE** : Especifica la codificación de los datos obtenidos de la forma. Este atributo únicamente se puede usar siempre y cuando **METHOD** tiene como valor **POST**<sup>50</sup>.

Ahora bien, una forma puede proporcionar diferentes elementos para la captura de datos, tales como cajas de texto, áreas de texto, cajas combo, menús desplegables, botones, imágenes, entre otros.

La etiqueta `<INPUT>` proporciona algunos de estos elementos. La sintaxis general es la siguiente :

```
<INPUT      {TYPE="tipo_de_elemento"}    {NAME="nombre"}    {SIZE=###}  
[VALUE="valor"] [CHECKED]>
```

El atributo **TYPE** especifica el tipo de elemento que se insertará en la forma. Los atributos que se puedan emplear van a variar de acuerdo con el tipo seleccionado. Los tipos de elemento son los siguientes :

- **TEXT**: Coloca una caja de texto para permitir a los usuario escribir en ella. El tamaño de dicha caja se especifica con el atributo **SIZE**, el número que se asigne será el número de letras que tendrá de longitud la caja. Si se omite el atributo **SIZE**, el tamaño será de 20. El texto que el usuario coloque en esta caja será procesado por el CGI declarado en **ACTION**. Si el atributo **VALUE** tiene asociada alguna cadena de caracteres, cuando se accesa a la forma por primera vez tendrá inicialmente dicho texto.

---

<sup>50</sup> Para mayor información de esta opción se recomienda revisar el documento titulado "Mosaic for X version 2.0 Fill-Out-Form-Support" en la dirección :  
<http://www.nesa.uiuc.edu/SDG/Software/Docs/ fill-out-forms/overview.html>

## Anexo B - CGI (Common Gateway Interface)

- **CHECKBOX:** Presenta un cuadro de selección que puede ser activado o desactivado. Se pueden utilizar varios elementos de este tipo para consultar diferentes alternativas (por ejemplo, se puede pedir los ingredientes de una Pizza ). Si el atributo *VALUE* tiene asociado algún valor y dicho elemento fue seleccionado por el usuario, al momento de enviar la información a procesar este elemento tendrá asociado ese valor. Si se desea que al iniciar uno o varios elementos de este tipo estén seleccionados, se utiliza el atributo *CHECKED* para cada uno de ellos.
- **RADIO:** Similar a *CHECKBOX* sólo que presenta círculos de selección en lugar de cuadros. Otra diferencia es que se pueden hacer grupos de estos elementos, para que sólo uno de ellos pueda ser seleccionado a la vez (Por ejemplo, se pueden utilizar para preguntar el estado civil de las personas). Se puede utilizar el atributo *VALUE* con los mismos efectos que en *CHECKBOX* al igual que con el atributo *CHECKED*.
- **SUBMIT:** Coloca un botón en la forma. Si se activa dicho botón la información obtenida en la forma se envía para ser procesada por el programa CGI especificado en *ACTION*. El botón tiene el título de *SUBMIT*, si se desea cambiar ese título se utiliza el atributo *VALUE* y se especifica un nuevo nombre. Éste botón es un elemento indispensable en una forma, ya que con él se procesa la información proporcionada por el usuario.
- **RESET:** Coloca un botón el cual "limpia" la forma. Esto es, devuelve a cada elemento el valor inicial con el que fue declarado. El botón tiene el título de *RESET*, y al igual que el botón *SUBMIT* hay cambiar el valor del atributo *VALUE* para especifica un nuevo nombre.
- **HIDDEN:** Este elemento no presenta nada. Se utiliza para enviar información que no se desea cambiar. Necesita los atributos *NAME* y *VALUE*.

En todos los casos el *NAME* sirve para que el programa CGI identifique cada uno de los elementos que llenó el usuario. Por ejemplo, si una caja de texto tiene asociado el nombre de "Nombre", el programa CGI identificará la variable Nombre con el valor con la cual le fue enviada. En otras palabras, es una referencia para los diferentes valores que se envía desde una forma.

Para colocar menús desplegables, se utiliza las etiquetas `<SELECT>` y `</SELECT>`. La sintaxis de este elemento es la siguiente :

```
<SELECT NAME="nombre">
  <OPTION VALUE="valor1" [SELECTED]> Nombre de la opción 1
  <OPTION VALUE="valor2" [SELECTED]> Nombre de la opción 2
  .....
  <OPTION VALUE="valorn" [SELECTED]> Nombre de la opción n
</SELECT>
```

## Anexo B - CGI (Common Gateway Interface)

Los atributos *NAME* y *VALUE* son los mismos que con el elemento *INPUT*. Si se desea que un elemento en particular esté seleccionado se utiliza el atributo *SELECTED* en dicha opción, de otra forma el primer elemento será seleccionado.

Otro elemento que puede ir en una forma son las áreas de texto. Un área de texto es un recuadro de varios renglones y columnas para la captura de texto largo. Se utiliza para obtener comentarios, sugerencias, etc. El texto capturado por éste elemento tiene actualmente un máximo de 256 caracteres actualmente. La sintaxis es la siguiente :

```
<TEXTAREA ROWS=### COLS=###>  
  Texto inicial  
</TEXTAREA>
```

El atributo *ROWS* indica el número de renglones del área de texto, mientras que *COLS* indica el número de columnas. El texto que se encuentre entre *<TEXTAREA>* y *</TEXTAREA>* servirá de texto inicial para este elemento.

Para enviar la información de cada uno de estos elementos, se tienen dos opciones : *GET* y *POST*. Si no se define algún método entonces se usa el método *GET*. Este envía la información como parte del URL del CGI definido en el atributo *ACTION* de la etiqueta *<FORM>*. El método *POST*, por su parte, establece una comunicación directa con el programa CGI al que hace referencia la forma, sin presentar los datos en el URL del programa.

Para mayor referencia sobre la implementación de CGIs se recomienda al lector los siguientes materiales escritos :

- El libro de CGI  
William E. Weinman  
Prentice Hall
- Building Internet Database Servers With CGI  
Jeff Rowe  
Prentice Hall  
1996

y los documentos en línea :

*Common Gateway Interface* : <http://hoohoo.ncsa.uiuc.edu/docs/cgi/intro.html>

*Decoding FORMs with CGI* : <http://hoohoo.ncsa.uiuc.edu/docs/cgi/forms.html>

GLOSARIO DE  
TÉRMINOS DE  
*INTERNET*

## Glosario de términos de *Internet*<sup>51</sup>

### **Ancho de banda**

Es la cantidad de información se pueden enviar a través de una conexión. Usualmente dicha cantidad se mide en bits por segundo. Una página completa de texto en Español tiene aproximadamente 16,000. Un módem rápido puede mover cerca de 15,000 bits por segundo. Una secuencia de pantalla completa de video puede requerir aproximadamente 10,000,000 bits por segundo, dependiendo de la compresión que tenga.

### **Applet**

Programa hecho con Java que puede ser incluido en una página HTML. Los Applets difieren de los programas Java en que no tienen permitido un completo control sobre ciertos recursos de la computadora local, tal como archivos y periféricos (módems, impresoras, etc.), y tienen restringida la comunicación con otras computadoras a través de una red. La regla actual es que un Applet sólo puede mantener la conexión con la computadora de la cual fue enviado.

### **Archie**

Un herramienta (*software*) para buscar archivos almacenados en un servidor FTP. Es necesario conocer el nombre exacto del archivo o una parte de él.

### **ARP**

(*Address Resolution Protocol*) – Convierte las direcciones IP de 32 bits en direcciones físicas de red (por ejemplo, direcciones Ethernet de 48 bits).

### **ARPANet**

(*Advanced Research Projects Agency NetWork*) – El precursor de *Internet*. Proyecto desarrollado a finales de los 60's y principios de los 70's por el departamento de Defensa de los E.U. como un experimento sobre redes de área amplia que pudiera sobrevivir a un ataque nuclear.

### **ASCII**

(*American Standard Code for Information Interchange*) – Este es el estándar mundial de facto para el código numérico usado por las computadoras en la representación de todas las letras Latinas mayúsculas y minúsculas, números, puntuación, etc. Existen 128 códigos de caracteres ASCII cada uno de ellos puede ser representado por un número binario del 00000000 al 11111111.

### **Backbone**

Línea de alta velocidad o serie de conexiones que forman el camino principal en una red. El término es relativo desde un *backbone* en una red pequeña, hasta aquel que pudiera ser mucho más grande que varias líneas en una red grande.

---

<sup>51</sup> Basado en el documento en línea titulado "ILC Glossary of Internet Terms": <http://www.matisse.net/files/glossary.html>

## Glosario de términos de Internet

### **Baud**

El baud es el número de veces por segundo que la señal portadora cambia. - por ejemplo un módem de 1200 bits por segundo actualmente corre a 300 bauds, pero mueve 4 bits por baud ( $4 \times 300 = 1200$  bits por segundo) - Es comúnmente utilizado en los módems para referirse a la cantidad de bits por segundo que pueden enviar o recibir.

### **Bit**

(*Binary digit*) -- Un dígito en base 2, sólo puede tener uno de dos valores: 0 ó 1 en otras palabras, tanto el 1 como el 0. La unidad más pequeña de información en una computadora

### **bps**

(Bits - por - segundo) – Medida de cuán rápido se mueve la información de un lugar a otro. Un módem de 28.8 puede mover 28,800 bits por segundo.

### **Browser o Navegador**

Un programa cliente (*software*) que se usa para desplegar varios tipos de documentos en la Internet. Generalmente este término se utiliza para hacer referencia a un Cliente de Web.

### **Byte**

Conjunto de Bits que representan un simple carácter. Usualmente hay 8 Bits en un Byte, algunas veces más, dependiendo en como se defina la medida.

### **CGI**

(*Common Gateway Interface*) – Conjunto de reglas que describen como se comunica un servidor Web con un programa en la misma u otra máquina, y como dicho programa (el "programa CGI") se comunica con el servidor Web. Cualquier programa puede ser un programa CGI si maneja la entrada y salidas de datos de acuerdo al estándar CGI.

### **cgi-bin**

Es el nombre más común de directorio de un servidor Web en el cual los programas CGI son almacenados.

### **Ciberespacio o Cyberspace**

Término originado por el autor William Gibson en su novela Neuromancer. La palabra Cyberspace (Ciberespacio) actualmente se usa para describir en su totalidad al rango de fuentes de información disponibles en las redes de computadoras.

### **Cliente**

Programa que se usa para contactar y obtener datos de un programa Servidor en otra máquina, que frecuentemente esta a gran distancia. Cada programa cliente se diseña para trabajar con uno o más tipos de programas Servidores, y cada Servidor requiere de una clase específica de Cliente. Un Navegador de Web es una clase específica de cliente.

### **Correo Electrónico**

Ver E-mail.

***Datagramas (o paquetes de información)***

Se denomina de esta forma a los bloques de datos en los que se ha dividido la información original para ser enviada a través de una red.

***Domain Name***

Nombre único que define un sitio en *Internet*. *Domain Name*, siempre tiene 2 o más partes, separadas por puntos. La parte de la izquierda es la más específica, y la parte de la derecha es más general. Una máquina cualquiera puede tener más de un *Domain Name*, pero para un *Domain Name* dado sólo le corresponde una máquina. Por ejemplo, los siguientes *Domain Name*

matisse.net  
mail.matisse.net  
workshop.matisse.net

pueden todos ellos hacer referencia a la misma máquina, pero cada dominio no puede referirse a más de un equipo.

***Dirección física***

Es la dirección que los componentes de red asignan a un equipo.

***E-mail***

(*Electronic Mail*) – Son mensajes, usualmente texto, enviados de una persona a otra vía computadoras. Estos mensajes pueden ser enviados automáticamente a un gran número de direcciones (Listas de correo).

***Espejo***

Generalmente en cómputo decir "crear un espejo", es para mantener una copia exacta de algo. Probablemente el uso más común de este término en *Internet* se refiere a los sitios "espejo", los cuales son sitios *Web* o *FTP* que mantienen copias exactas del material que originalmente se tiene en otra locación; este material usualmente se encuentra en orden para proveer un acceso más extendido.

***Ethernet***

Un método muy común de disposición de redes LAN. *Ethernet* maneja cerca de 10,000,000 bits por segundo y puede ser usado con muchas clases de computadoras.

***FAQ***

(*Frequently Asked Question*) – Son documentos que listan y responden las preguntas más comunes sobre un tema en particular. Hay miles de *FAQs* sobre temas tan diversos como Cuidado de mascotas y Criptografía.

***Finger***

Una herramienta (programa) para *Internet* que permite localizar personas en otros sitios en la *Internet*. *Finger* es algunas veces usado para tener acceso a información no personal, pero el uso más común es para ver si una persona tiene una cuenta en un sitio en la *Internet* en particular.

### **Fire Wall**

Una combinación de *software* y *hardware* que separa una red LAN en dos o más partes con propósitos de seguridad.

### **FTP**

(*File Transfer Protocol*) – Un método muy común de transferir archivos entre dos máquinas en la Internet. FTP es una forma especial para ingresar a un equipo en *Internet* con el propósito de obtener y/o enviar archivos. Existen muchos sitios en *Internet* que han establecido almacenes de *software* de acceso público que puede ser obtenido usando FTP, proporcionando la clave anonymous como nombre de la cuenta, dichos sitios son nombrados Servidores FTP anónimos.

### **Gateway**

El significado técnico es: un *hardware* o *software* apropiado que traduce entre dos protocolos similares, por ejemplo, *Prodigy* tiene un *gateway* que traduce entre su formato propio e interno de correo electrónico y el utilizado por *Internet*.

### **Gopher**

Método ampliamente exitoso para crear menús de material disponible en *Internet*. Gopher es un servicio Cliente/Servidor, el cual requiere que el usuario tenga un programa Cliente para Gopher. Aunque Gopher se extendió rápidamente a través del globo en un par de años, está siendo sustituido ampliamente por el hipertexto (WWW - *World Wide Web*).

### **Gigabyte**

1000 Kilobytes

### **Home Page (o Homepage)**

Tiene muchos significados. Originalmente se denominó así a la página Web que en un Navegador se configuraba para el inicio de una sesión. El significado más aceptado se refiere al de la página principal para un negocio, organización persona o simplemente la página principal de un sitio.

### **Host**

Cualquier computadora en una red que tiene depositado servicios disponibles para otras computadoras en la red. Es común tener una máquina que provea varios servicios, tales como WWW y USENET.

### **Hipertexto**

Generalmente, cualquier texto que contiene ligas a otros documentos. Palabras o frases en el documento que al ser seleccionadas por el usuario provocan que otro documento sea recuperado y desplegado.

### **HTML**

(*HyperText Markup Language*) – Lenguaje usado para crear documentos con hipertexto usados en el *World Wide Web*. El HTML se parece mucho a los viejos códigos de edición de texto, donde se rodea un bloque de texto con códigos que indican cómo debe de aparecer, además, en HTML se puede especificar que un bloque de texto o palabra, este ligada a otro archivo en *Internet*. Los archivos HTML por lo general se visualizan usando un programa cliente de WWW, tal como Netscape o Mosaic.

### **HTTP**

(*HyperText Transport Protocol*) – Es el protocolo para mover archivos de hipertexto a través de la red *Internet*. Requiere un programa Cliente de HTTP en un extremo, y un programa Servidor de HTTP del otro lado. HTTP es el protocolo más importante usado en *World Wide Web* (WWW).

### **Internet**

(I mayúscula). La vasta colección de redes interconectadas que usan los protocolos TCP/IP y que evolucionó de la ARPANet de finales de los 60s y principios de los 70s. La red *Internet*, hoy en día (Julio de 1995), conecta 60,000 redes independientes en una vasta *internet*.

### **internet**

(i minúscula). En el momento en que dos o más redes se conectan entre sí, se tiene una *internet* - como en *inter-relación*, *inter-estatal*, etc.

### **Intranet**

Red privada dentro de una compañía u organización que usa los mismos tipos de programas que se encuentran en la *Internet*, pero sólo internamente.

Como *Internet* se ha hecho muy popular, muchas de sus herramientas están siendo utilizadas en redes privadas, por ejemplo, muchas compañías tienen servidores Web que están disponibles únicamente para sus empleados.

Cabe hacer notar que una *Intranet* puede no ser una *internet*, es decir puede ser simplemente una red.

### **IP**

(*Internet Protocol*) – Protocolo que dirige los datos a través de redes separadas vinculadas mediante *Routers* para formar la *Internet* o una *Intranet*.

### **ISDN**

(*Integrated Services Digital Network*) – Básicamente una manera de mover más datos sobre las líneas de teléfonos existentes. ISDN están siendo disponible rápidamente en los E.U.

### **Java**

Lenguaje de programación orientado a objetos desarrollado por Sun Microsystems. Fue diseñado específicamente para escribir programas que puedan ser obtenidos a través de la *Internet* sin peligro de que causen daño al equipo y/o archivos locales.

### **JDK**

(*Java Development Kit*) – Paquete para desarrollo de software creado por Sun Microsystems que implementa el conjunto básico de herramientas necesarias para escribir, probar y depurar aplicaciones y Applets Java.

### **Kilobyte**

Aproximadamente mil bytes. Con precisión, 1024 (2<sup>10</sup>) bytes.

## Glosario de términos de *Internet*

### **LAN**

(*Local Area Network*) – Red de computadoras limitada a un área inmediata, siendo usualmente el mismo edificio o piso del mismo.

### **Lista de correo**

Sistema generalmente automatizado que permite a las personas enviar *E-mail* a una dirección. En este sistema, los mensajes se copian y envían al resto de los suscriptores de la lista. De esta manera, quienes tienen claves de correo pueden participar juntos en discusiones.

### **Login**

Sustantivo o verbo. Sustantivo : Nombre de la cuenta usada para obtener el acceso a un sistema de cómputo. No es secreto, opuesto a *Password*.  
Verbo : El acto de acceder a un sistema de cómputo.

### **Megabyte**

Un millón de bytes. Mil kilobytes.

### **Módem**

(*MOdulator, DEModulator*) – Dispositivo que conecta a una computadora con la línea telefónica, permitiendo a la primera comunicarse con otras computadoras a través del sistema telefónico.

### **Mosaic**

El primer Navegador de WWW que estuvo disponible para Macintosh, Windows y UNIX, todos con la misma interfaz gráfica. Fue desarrollado en el Centro Nacional de Aplicaciones de Supercómputo (NCSA – *National Center for Supercomputing Applications*). Con *Mosaic* realmente comenzó la popularización del *Web*. El código fuente ha sido comercializado por varias compañías y hay mucho otros programas tan buenos o mejores como este.

### **Netscape**

Navegador de Web y nombre de la misma compañía. El Navegador Netscape™ se basó originalmente en el programa *Mosaic*. El autor principal de Netscape, Mark Andreesen fue contratado por Jim Clark, y ambos fundaron una compañía llamada Mosaic Communications la cual pronto cambió su nombre a Netscape Communications Corporation.

### **NIC**

(*Networked Information Center*) – Generalmente, es una oficina que maneja información por una red. El más famoso NIC en la Internet es el InterNIC, el cual se encuentra localizado donde se registran los nuevos *Domain Names*.

### **Nodo**

Cualquier computadora conectada a una red.

### **Nombre de dominio**

Ver *Domain Name*.

### Número IP o Dirección IP

Un número único consistente de 4 partes separadas por comas. Algunas veces se le refiere como notación de puntos. Un ejemplo de número IP es el siguiente :

132.248 71.5

Cada equipo que este conectado a la Internet tiene un número IP único – Si una máquina no tiene un número IP, realmente no esta en la Internet.

### Paquetes de información

Ver Datagramas

### Password

Código usado para acceder a un sistema cerrado. Buenos passwords contienen caracteres alfanuméricos en una combinación no sencilla, tal como *virtue7*. Un buen password puede ser :

HotS1-6

### POP

(*Point Of Presence*, también *Post Office Protocol*) – Dos significados usados comúnmente : Punto de presencia y Protocolo de oficina postal. Un Punto de presencia usualmente significa una ciudad o comunidad donde una red puede ser conectada, frecuentemente, por líneas telefónicas. El segundo significado, Protocolo de oficina postal se refiere a la forma en que el software e-mail obtiene el correo de un servidor.

### Port

Tres significados. El primero y más general, un lugar donde la información va de y hacia una computadora, o ambos. Por ejemplo, el puerto serial de una computadora personal es donde un módem debe ser colocado

En Internet, el puerto a menudo se refiere a un número del URL, que aparece a la derecha del *Domain Name*, después de los dos puntos (:). Por ejemplo, en el siguiente URL el puerto es el número 7000 :

`gopher://peg.cwis.uci.edu:7000/`

Finalmente, un puerto también se refiere a la transferencia de un programa de un tipo de computadora a otro. por ejemplo para transferir un programa de Windows tal que correrá en una Macintosh.

### PPP

(*Point to Point Protocol*) – Más conocido como un protocolo que permite a una computadora usar la línea telefónica y un módem para utilizar conexiones TCP/IP de tal manera que esté en la Internet

### RARP

(*Reverse Address Resolutio Protocol*) – Convierte direcciones físicas de red en direcciones IP.

## Glosario de términos de *Internet*

### **Router**

Una computadora de propósito particular (o programa) que controla la conexión entre 2 o más redes. Los *Routers* consumen todo su tiempo buscando las direcciones destino de los paquetes de información a través de ellos y deciden por cuál ruta los enviará.

### **Servidor**

Es una computadora, o un programa, que provee un tipo especial de servicio para un programa *Software* corriendo en otras computadoras. El término puede referirse a un programa en particular, tal como un servidor *WWW*, o la máquina sobre la cual el programa está ejecutándose.

### **SLIP**

(*Serial Line Internet Protocol*) – Un estándar para el uso de la línea telefónica (una línea serial) y un módem para conectar una computadora como un sitio real de *Internet*. *SLIP* está siendo gradualmente reemplazado por *PPP*.

### **SQL**

(*Structured Query Language*) – Un lenguaje de programación especializado para enviar consultas a bases de datos. Importantes compañías y pequeñas aplicaciones de bases de datos están usando *SQL*. Cada aplicación en particular tendrá su propia versión de *SQL* implementando características únicas para dicha aplicación, pero toda la especificación para *SQL* estará disponible.

### **TCP**

(*Transmission Control Protocol*) – Un protocolo orientado a conexiones que transmite información en flujos de bytes. La información se transmite en segmentos llamados paquetes de información. *TCP* es un protocolo "confiable" ya que utiliza la verificación de errores para constatar la integridad de la información y asegura que la información transmitida se ha recibido intacta.

### **TCP/IP**

(*Transmission Control Protocol/Internet Protocol*) – Es un conjunto de protocolos que definen la red *Internet*. Originalmente fue diseñado para el sistema operativo *UNIX*, los programas de *TCP/IP* ahora están disponibles para las principales clases de sistemas operativos. Para que un equipo esté realmente conectado a la *Internet*, debe contar con programas de *TCP/IP*.

### **Terabyte**

1000 gigabytes

### **Terminal**

Dispositivo que permite el envío de comandos a una computadora que se localice en otra parte. Básicamente, esto significa un teclado, un monitor y algo de circuitería. Generalmente se dispone de un programa en una computadora personal que emula una terminal física y permite al usuario escribir comandos a una computadora en otro lado.

### **UDP**

(*User Datagram Protocol*) – Protocolo libre de conexiones que transmite la información en paquetes de datos (datagramas). *UDP* es un protocolo "no confiable" ya que el transmisor no recibe información que indique si el datagrama fue en realidad recibido.

### **UNIX**

Sistema operativo de computadora. UNIX se diseñó para ser utilizado por muchas personas al mismo tiempo (esto es multi-usuario) y tiene TCP/IP interconstruido. Es el sistema operativo de mayor uso para servidores en la Internet.

### **URL**

(*Uniform Resource Locator*) – Ruta estándar para dar la dirección de una fuente de información en el *World Wide Web*, que es parte de la Internet. Un URL es algo como lo siguiente :

http://www.unam.mx/Temas  
o telnet://azteca.dgsca.unam.mx  
o ftp://ftp.ciencias.unam.mx/  
etc.

El uso más común de un URL es para obtener documentos HTML cuando se proporciona a un Navegador de WWW, tal como NetScape o Lynx

### **Veronica**

(*Very Easy Rodent Oriented Net-wide Index to Computerized Archives*) – Desarrollado por la Universidad de Nevada, Veronica es una base de datos de nombres de casi todos los títulos de menú de los servidores de Gopher constantemente actualizada. Esta base de datos puede ser revisada por la mayoría de los menús de Gopher.

### **WAIS**

(*Wide Area Information Servers*) – Un programa comercial que permite la indexación de una enorme cantidad de información, y dichos índices pueden ser consultados a través de las redes tal como la Internet.

### **Web o WWW**

(*World Wide Web*) – Dos significados. El primero, vagamente usado : toda la constelación de fuentes de información que pueden ser accedidas usando Gopher, FTP, HTTP, telnet, WAIS y algunas otras herramientas. La segunda, el universo de los servidores de hipertexto (servidores de HTTP) los cuales son ahora también servidores de texto, gráficas, archivos de sonido, etc. para ser mezclados en un solo documento.

# BIBLIOGRAFÍA

**Internet paso a paso - Hacia la autopista de la información**

**Gonzalo Ferreyra C.**

**Edit. COMPUTEC**

**México 1996**

**HTML & CGI UNLEASHED**

**John December, Mark Ginsburg**

**SAMS NET, 1a edición**

**E.U. - 1995**

**Conéctate al mundo de INTERNET**

**De Krol**

**O'Reilley & Associates, Inc. 2a edición**

**México - 1994**

**WebSite - Building your own WebSite**

**Susan B. Peck, Stephen Arrants**

**O'Reilly 1a edición**

**E.U. - Julio 1996**

**Teach yourself The Internet in a week**

**Neil Randall, Jhon December, Caro DeVrieze, otros**

**SAMS NET 2a edición**

**E.U. - 1995**

**WEB site Construction Kit for Windows 95**

**Scott Zimmermany, Christopher L. T. Brown**

**Prentice Hall 1a edición**

**E.U. - Diciembre 1996**

**Building Internet Database Servers With CGI**

**Jeff Rowe**

**Prentice Hall, 1a edición**

**E.U. - Diciembre 1996**

**El libro de CGI**

**William E. Weinman**

**Prentice Hall**

**México - Diciembre 1996**